WO2004091524

Publication Title:

RESPIRATORY VIRUS VACCINES

Abstract:

The invention provides vaccines and methods for preventing, treating, and diagnosing respiratory virus infection.

---

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 October 2004 (28.10.2004)

PCT

(10) International Publication Number
WO 2004/091524 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number: PCT/US2004/011425

(22) International Filing Date: 14 April 2004 (14.04.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/462,995 14 April 2003 (14.04.2003) US
60/515,010 28 October 2003 (28.10.2003) US

(71) Applicant *(for all designated States except US)*: **ACAMBIS INC.** [US/US]; 38 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MONATH, Thomas, P.** [US/US]; 21 Finn Road, Harvard, MA 01451 (US). **KLEANTHOUS, Harold** [GB/US]; 51 Russell's Way, Westford, MA 01886 (US).

(74) Agent: **MICHAUD, Susan, M.**; Clark & Elbing LLP, 101 Federal Street, Boston, MA 02110-1817 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: RESPIRATORY VIRUS VACCINES

(57) Abstract: The invention provides vaccines and methods for preventing, treating, and diagnosing respiratory virus infection.

WO 2004/091524 A2

PATENT
ATTORNEY DOCKET NO. 06132/080WO3

RESPIRATORY VIRUS VACCINES

Background of the Invention

Severe Acute Respiratory Syndrome (SARS) is a life-threatening respiratory illness that has recently been reported in Asia, North America, and Europe. SARS is thought to have originated in the Guangdong Province of China, and then to have been transported to Hong Kong by an infected healthcare worker who, when visiting Hong Kong, was hospitalized and died. SARS is thought to be transmissible in droplet form. Thus, it may be transmitted when an infected individual coughs or sneezes droplets into the air, and someone else breathes them in. SARS may also be transmitted more broadly through the air, or by the touching of objects that are contaminated. The illness usually begins with a fever, often accompanied by chills, headache, general discomfort, body aches, and/or mild respiratory symptoms. As the disease progresses, some patients develop a dry, non-productive cough. In addition, in some cases, the disease can progress to the point where mechanical ventilation is required to enable sufficient oxygen to enter a patient's bloodstream.

Viruses in the *Coronaviradae* family are characterized by a halo or crown-like (corona) appearance on their outer shell when viewed by microscopy. These viruses are a common cause of mild to moderate upper-respiratory illness in humans, and may account for up to one-third of cases of the common cold. Coronaviruses are also often found in animals, such as chickens, pigs, dogs, and cats, in which they can cause illnesses that range from diarrhea to respiratory infection. Further, coronaviruses have been found to survive in the environment for as long as three hours. It has been determined that a previously unrecognized coronavirus can be found in samples from patients with SARS.

Summary of the Invention

The invention provides vaccines for inducing an immune response to a human coronavirus that is the causative agent of Severe Acute Respiratory Syndrome (SARS) in a patient. These vaccines can include a spike protein and/or a nucleocapsid protein of the virus, or immunogenic fragments of either or both of these proteins, and a

pharmaceutically acceptable carrier or diluent. Specific examples of spike protein fragments that can be included in the vaccine compositions of the invention are those including the S1 domain, the S1 domain and the S2 domain, in the absence of the coiled coil region, and the S1 and S2 domains, including the coiled coil domain. Further, the spike protein (or fragment) can be present in the form of a monomer, a dimer, or a trimer.

Optionally, the vaccine compositions can also include an adjuvant, such as an adjuvant that stimulates a Th1-type immune response (e.g., an ISCOM, Ribi, DC-Chol, QS21, or MPL). Another example of an adjuvant that can be included in the vaccines of the invention is aluminum hydroxide (e.g., alum). In one example, the proteins of the vaccines of the invention include an amino acid sequence that is substantially identical to the sequence of SEQ ID NO:37 or SEQ ID NO:35, or immunogenic fragments thereof.

The invention also includes additional vaccines for inducing an immune response to human coronaviruses that cause SARS. These vaccines include vectors (e.g., viral vectors) containing a nucleic acid sequence encoding a spike protein or a nucleocapsid protein of the virus, or an immunogenic fragment thereof, and a pharmaceutically acceptable carrier or diluent. An example of a vector that can be used in such vaccines is a poxvirus, such as a Modified Vaccinia Ankara (MVA) vector. Another example of such a vector is adenovirus vectors.

The invention also provides methods for producing spike proteins or nucleocapsid proteins of human coronaviruses that cause SARS. These methods involve introducing into cells a vector that includes a nucleic acid sequence encoding the protein, under conditions in which the protein is expressed in the cells. These cells can be, for example, yeast cells, mammalian cells, insect, or bacterial cells.

The invention further provides methods of inducing an immune response to a human coronavirus that causes SARS in patients, by administration of the vaccines described above and elsewhere herein to the patients. The immune response can be prophylactic or therapeutic.

Also, the invention provides substantially pure spike proteins of human coronaviruses that cause SARS, or immunogenic fragments thereof. For example, such a protein can include a sequence that is substantially identical to or identical to the

sequence of SEQ ID NO:37, or a fragment thereof. The spike proteins and fragments of the invention can be in the form of monomers, dimers, or trimers.

The invention also includes isolated nucleic acid molecules encoding spike proteins of human coronaviruses that cause SARS. Such a nucleic acid molecule can include the sequence of SEQ ID NO:36, or a sequence that hybridizes to the complement of the sequence of SEQ ID NO:36 under highly stringent conditions. The invention also includes nucleic acid molecule probes that include sequences that hybridize to the sequence of SEQ ID NO:36 or the complement thereof under highly stringent conditions.

In addition, the invention provides substantially pure nucleocapsid proteins of human coronaviruses that cause SARS, or immunogenic fragments thereof. For example, such a protein can include a sequence that is substantially identical to or identical to the sequence of SEQ ID NO:35, or a fragment thereof.

The invention also includes isolated nucleic acid molecules encoding nucleocapsid proteins of human coronaviruses that cause SARS. Such a nucleic acid molecule can include the sequence of SEQ ID NO:34, or a sequence that hybridizes to the complement of the sequence of SEQ ID NO:34 under highly stringent conditions. The invention also includes nucleic acid molecule probes that include sequences that hybridize to the sequence of SEQ ID NO:34 or the complement thereof under highly stringent conditions.

Further, the invention includes antibodies (e.g., monoclonal, monospecific, and polyclonal antibodies) that specifically bind to spike proteins or nucleocapsid proteins of human coronaviruses that cause SARS. These antibodies can be used in passive immunization methods, as described elsewhere herein.

By "polypeptide" or "polypeptide fragment" is meant a chain of two or more (e.g., 10, 15, 20, 30, 50, 100, or 200, or more) amino acids, regardless of any post-translational modification (e.g., glycosylation or phosphorylation), constituting all or part of a naturally or non-naturally occurring polypeptide. By "post-translational modification" is meant any change to a polypeptide or polypeptide fragment during or after synthesis. Post-translational modifications can be produced naturally (such as

during synthesis within a cell) or generated artificially (such as by recombinant or chemical means). A "protein" can be made up of one or more polypeptides.

By "spike protein" or "spike polypeptide" is meant a polypeptide that has at least 45%, preferably at least 60%, more preferably at least 75%, 80%, or 85%, and most preferably at least 90%, 95%, 99%, or 100% amino acid sequence identity to the sequence of SEQ ID NO:37. These proteins and polypeptides (or fragments thereof, as well as corresponding nucleic acid molecules) can be used in vaccines as described herein, as well as for markers of infection by human coronaviruses that cause SARS.

By "SARS nucleocapsid protein" or "SARS nucleocapsid polypeptide" is meant a polypeptide that has at least 45%, preferably at least 60%, more preferably at least 75%, 80%, or 85%, and most preferably at least 90%, 95%, 99%, or 100% amino acid sequence identity to the sequence of SEQ ID NO:35. These proteins and polypeptides (or fragments thereof, as well as corresponding nucleic acid molecules) can be used in vaccines as described herein, as well as for markers of infection by human coronaviruses that cause SARS.

Useful polypeptide derivatives, e.g., polypeptide fragments, can be designed using computer-assisted analysis of amino acid sequences in order to identify sites in protein antigens having potential as surface-exposed, antigenic regions (see, e.g., Hughes et al., Infect. Immun. 60(9):3497, 1992). For example, the Laser Gene Program from DNA Star can be used to obtain hydrophilicity, antigenic index, and intensity index plots for the polypeptides of the invention. This program can also be used to obtain information about homologies of the polypeptides with known protein motifs. One skilled in the art can readily use the information provided in such plots to select peptide fragments for use as vaccine antigens. For example, fragments spanning regions of the plots in which the antigenic index is relatively high can be selected. Fragments spanning regions in which both the antigenic index and the intensity plots are relatively high can also be selected, as well as fragments containing conserved sequences, particularly hydrophilic conserved sequences.

By a "spike nucleic acid molecule" is meant a nucleic acid molecule, such as a genomic DNA, cDNA, or RNA (e.g., mRNA) molecule, that encodes a spike protein

(e.g., a protein encoded by SEQ ID NO:36), a spike polypeptide, or a portion thereof, as defined above.

By a "SARS nucleocapsid protein nucleic acid molecule" is meant a nucleic acid molecule, such as a genomic DNA, cDNA, or RNA (e.g., mRNA) molecule, that encodes a spike protein (e.g., a protein encoded by SEQ ID NO:34), a nucleocapsid polypeptide, or a portion thereof, as defined above.

The term "identity" is used herein to describe the relationship of the sequence of a particular nucleic acid molecule or polypeptide to the sequence of a reference molecule of the same type. For example, if a polypeptide or a nucleic acid molecule has the same amino acid or nucleotide residue at a given position, compared to a reference molecule to which it is aligned, there is said to be "identity" at that position. The level of sequence identity of a nucleic acid molecule or a polypeptide to a reference molecule is typically measured using sequence analysis software with the default parameters specified therein, such as the introduction of gaps to achieve an optimal alignment (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705, BLAST, or PILEUP/PRETTYBOX programs). These software programs match identical or similar sequences by assigning degrees of identity to various substitutions, deletions, or other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine, valine, isoleucine, and leucine; aspartic acid, glutamic acid, asparagine, and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine.

The sequence of a nucleic acid molecule or polypeptide is said to be "substantially identical" to that of a reference molecule if it exhibits at least 51%, preferably at least 55%, 60%, or 65%, and most preferably 75%, 85%, 90%, or 95% identity to the sequence of the reference molecule. For polypeptides, the length of comparison sequences is at least 16 amino acids, preferably at least 20 amino acids, more preferably at least 25 amino acids, and most preferably at least 35 amino acids. For nucleic acid molecules, the length of comparison sequences is at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most

preferably at least 110 nucleotides. Of course, for polypeptides and nucleic acid molecules, the length of comparison can be any length up to and including full length.

By "probe" or "primer" is meant a single-stranded DNA or RNA molecule of defined sequence that can base pair to a second DNA or RNA molecule that contains a complementary sequence (a "target"). The stability of the resulting hybrid depends upon the extent of the base pairing that occurs. This stability is affected by parameters such as the degree of complementarity between the probe and target molecule, and the degree of stringency of the hybridization conditions. The degree of hybridization stringency is affected by parameters such as the temperature, salt concentration, and concentration of organic molecules, such as formamide, and is determined by methods that are well known to those skilled in the art. Probes or primers specific for spike or nucleocapsid nucleic acid molecules, preferably, have greater than 45% sequence identity, more preferably at least 55-75% sequence identity, still more preferably at least 75-85% sequence identity, yet more preferably at least 85-99% sequence identity, and most preferably 100% sequence identity to the sequences of genes encoding spike or nucleocapsid proteins of a SARS-causing human coronavirus (SEQ ID NOs:36 and 34, respectively). Probes can be detectably labeled, either radioactively or non-radioactively, by methods that are well known to those skilled in the art. Probes can be used for methods involving nucleic acid hybridization, such as nucleic acid sequencing, nucleic acid amplification by the polymerase chain reaction, single stranded conformational polymorphism (SSCP) analysis, restriction fragment polymorphism (RFLP) analysis, Southern hybridization, northern hybridization, in situ hybridization, electrophoretic mobility shift assay (EMSA), and other methods that are well known to those skilled in the art.

A molecule, e.g., an oligonucleotide probe or primer, a gene or fragment thereof, a cDNA molecule, a polypeptide, or an antibody, can be said to be "detectably-labeled" if it is marked in such a way that its presence can be directly identified in a sample. Methods for detectably labeling molecules are well known in the art and include, without limitation, radioactive labeling (e.g., with an isotope, such as ^{32}P or ^{35}S) and nonradioactive labeling (e.g., with a fluorescent label, such as fluorescein).

By a "substantially pure polypeptide" is meant a polypeptide (or a fragment thereof) that has been separated from proteins and organic molecules that naturally accompany it. Typically, a polypeptide is substantially pure when it is at least 60%, by weight, free from the proteins and naturally occurring organic molecules with which it is naturally associated. Preferably, the polypeptide is a spike or nucleocapsid polypeptide that is at least 75%, 80%, or 85%, more preferably at least 90%, and most preferably at least 99%, by weight, pure. A substantially pure spike or nucleocapsid polypeptide can be obtained, for example, by extraction from a natural source, by expression of a recombinant nucleic acid molecule encoding a spike or nucleocapsid polypeptide, or by chemical synthesis. Purity can be measured by any appropriate method, e.g., by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

A polypeptide is substantially free of naturally associated components when it is separated from those proteins and organic molecules that accompany it in its natural state. Thus, a protein that is chemically synthesized or produced in a cellular system that is different from the cell in which it is naturally produced is substantially free from its naturally associated components. Accordingly, substantially pure polypeptides not only include those that are derived from coronaviruses, but also those synthesized in yeast systems. insect systems, mammalian systems, *E. coli*, other prokaryotes, or in other such systems (see below).

By "isolated nucleic acid molecule" is meant a nucleic acid molecule that is removed from the environment in which it naturally occurs. For example, a naturally-occurring nucleic acid molecule present in the genome of cell or as part of a gene bank is not isolated, but the same molecule, separated from the remaining part of the genome, as a result of, e.g., a cloning event (amplification), is "isolated." Typically, an isolated nucleic acid molecule is free from nucleic acid regions (e.g., coding regions) with which it is immediately contiguous, at the 5' or 3' ends, in the naturally occurring genome. Such isolated nucleic acid molecules can be part of a vector or a composition and still be isolated, as such a vector or composition is not part of its natural environment.

An antibody is said to "specifically bind" to a polypeptide if it recognizes and binds to the polypeptide (e.g., a spike or nucleocapsid polypeptide), but does not

substantially recognize and bind to other molecules (e.g., non-spike-related or non-nucleocapsid-related polypeptides) in a sample, e.g., a biological sample, which naturally includes the polypeptide. Antibodies that specifically bind to the spike or nucleocapsid proteins of human coronaviruses causing SARS are also included in the invention.

By "high stringency conditions" is meant conditions that allow hybridization comparable with the hybridization that occurs using a DNA probe of at least 100, e.g., 200, 350, or 500, nucleotides in length, in a buffer containing 0.5 M $NaHPO_4$, pH 7.2, 7% SDS, 1 mM EDTA, and 1% BSA (fraction V), at a temperature of 65°C, or a buffer containing 48% formamide, 4.8 x SSC, 0.2 M Tris-Cl, pH 7.6, 1 x Denhardt's solution, 10% dextran sulfate, and 0.1% SDS, at a temperature of 42°C. (These are typical conditions for high stringency northern or Southern hybridizations.) High stringency hybridization is also relied upon for the success of numerous techniques routinely performed by molecular biologists, such as high stringency PCR, DNA sequencing, single strand conformational polymorphism analysis, and in situ hybridization. In contrast to northern and Southern hybridizations, these techniques are usually performed with relatively short probes (e.g., usually 16 nucleotides or longer for PCR or sequencing, and 40 nucleotides or longer for in situ hybridization). The high stringency conditions used in these techniques are well known to those skilled in the art of molecular biology, and examples of them can be found, for example, in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, 1998, which is hereby incorporated by reference.

The invention provides several advantages. First, the invention provides approaches to preventing, treating, diagnosing a severe, life-threatening disease that has recently appeared in outbreaks around the world, in a short period of time. Further, the invention provides expression and vector systems that can be used to achieve high levels of expression and efficient delivery of SARS proteins, respectively.

Other features and advantages of the invention will be apparent from the following detailed description, the drawings, and the claims.

Brief Description of the Drawings

Figures 1-36 are schematic illustrations of constructs used in the expression of SARS spike proteins in *Pichia pastoris*, CHO cells, and Drosophila S2 cells.

In particular, Figure 1 provides the deduced amino acid sequence of pPICZ alpha 1190 clone P5-12 (SEQ ID NO:1); Figure 2 provides a linear map of the construct, including the AOX promoter, alpha signal sequence, spike amino acids 14-1190, and the AOX terminator sequence; Figure 3 provides a circular map of the construct; and Figure 4 provides the nucleotide sequence of this clone, based on the linear map (SEQ ID NO:2).

Figure 5 provides the deduced amino acid sequence of pPICZ alpha 709 clone P1-2 (SEQ ID NO:3); Figure 6 provides a linear map of the construct, including the AOX promoter, alpha signal sequence, spike amino acids 14-709, and the AOX terminator sequence; and Figure 7 provides the nucleotide sequence of the clone, based on the linear map (SEQ ID NO:4).

Figure 8 provides the deduced amino acid sequence of pPICZ alpha 719 clone P1-2 (SEQ ID NO:5); Figure 9 provides a linear map of the construct, including the AOX promoter, alpha signal sequence, spike amino acids 14-719, and the AOX terminator sequence: and Figure 10 provides the nucleotide sequence of the clone, based on the linear map (SEQ ID NO:6).

Figure 11 provides the deduced amino acid sequence of pPICZ alpha 883 clone P3-10 (SEQ ID NO:7); Figure 12 provides a linear map of the construct, including the AOX promoter, alpha signal sequence, spike amino acids 14-883, and the AOX terminator sequence; and Figure 13 provides the nucleotide sequence of the clone, based on the linear map (SEQ ID NO:8).

Figure 14 provides the deduced amino acid sequence of pPICZ alpha 883m clone P3-10 (SEQ ID NO:9); Figure 15 provides a linear map of the construct, including the AOX promoter, alpha signal sequence, spike amino acids 14-883, and the AOX terminator sequence; and Figure 16 provides the nucleotide sequence of the clone, based on the linear map (SEQ ID NO:10).

Figure 17 provides a circular map of pGAPZ alpha 1190 clone G5-14; Figure 18 provides the deduced amino acid sequence of the clone (SEQ ID NO:11); Figure 19 provides a linear map of the construct, including the GAP promoter, alpha signal sequence, spike amino acids 14-1190, and the AOX terminator sequence; and Figure 20 provides the nucleotide sequence of the clone (SEQ ID NO:12).

Figure 21 provides a linear map of pGAPZ alpha 709 clone G1-8, including the GAP promoter, alpha signal sequence, spike amino acids 14-709, and the AOX terminator sequence; Figure 22 provides the nucleotide sequence of the clone (SEQ ID NO:13); and Figure 23 provides the deduced amino acid sequence of the clone (SEQ ID NO:14).

Figure 24 provides the deduced amino acid sequence of pGAPZ alpha 719 clone G1-8 (SEQ ID NO:15); Figure 25 provides a linear map of the construct, including the GAP promoter, alpha signal sequence, spike amino acids 14-719, and the AOX terminator sequence; and Figure 26 provides the nucleotide sequence of the clone (SEQ ID NO:16).

Figure 27 provides the deduced amino acid sequence of pGAPZ alpha 883 clone G3-7 (SEQ ID NO:17); Figure 28 provides a linear map of the construct, including the GAP promoter, alpha signal sequence, spike amino acids 14-883, and the AOX terminator sequence; and Figure 29 provides the nucleotide sequence of the clone (SEQ ID NO:18).

Figure 30 provides the deduced amino acid sequence of pGAPZ alpha 883m clone G3-7 (SEQ ID NO:19); Figure 31 provides a linear map of the construct, including the GAP promoter, alpha signal sequence, spike amino acids 14-883, and the AOX terminator sequence; and Figure 32 provides the nucleotide sequence of the clone (SEQ ID NO:20).

Figure 33 provides a linear map of pMT-Spike 1190 and the nucleotide (SEQ ID NO:21) and amino acid (SEQ ID NO:22) sequences of this construct.

Figure 34 provides a linear map of pMT-Spike 719 and the nucleotide (SEQ ID NO:23) and amino acid (SEQ ID NO:24) sequences of this construct.

Figure 35 provides a linear map of pMT-Spike 883 and the nucleotide (SEQ ID NO:25) and amino acid (SEQ ID NO:26) sequences of this construct.

Figure 36 provides a linear map of pSec1190 and the nucleotide (SEQ ID NO:27) and amino acid (SEQ ID NO:28) sequences of this construct.

Figure 37 provides a linear map of pSec719 and the nucleotide (SEQ ID NO:29) and amino acid (SEQ ID NO:30) sequences of this construct.

Figure 38 provides a linear map of pSec883 and the nucleotide (SEQ ID NO:31) and amino acid (SEQ ID NO:32) sequences of this construct.

Figure 39 is a schematic representation of the structure of SARS S protein and target antigenic domains selected for expression.

Figure 40 is a schematic representation of approaches described herein for obtaining S protein expression in the hosts *Pichia pastoris*, Drosophila S2 Schneider, and CHO cells.

Figure 41 is a schematic representation of a generalized strategy for constitutive (CHO) and inducible (S2) expression of recombinant spike protein.

Figure 42 shows PCR screening and Western blot analysis of transiently transfected S2 cells.

Figure 43 shows RT-PCR confirmation of mRNA synthesis of S protein candidates 719, 883, and 1190 in CHO cells.

Figure 44 is a schematic representation of a generalized strategy for expression of recombinant S protein in *Pichia pastoris*.

Figure 45 shows S gene specific PCR confirming integration into *Pichia pastoris*.

Figure 46 shows constitutive expression of the S protein in *Pichia pastoris*.

Figure 47 shows a scheme for fractionation of high molecular weight S glycoprotein, as well as analysis of the immunoreactivity of the high molecular weight complex.

Figure 48 shows a scheme for purification of high molecular weight S glycoprotein (1190), as well as immunoblot analysis of the purified material.

Figure 49 shows Anti-SARS-CoV (hyperimmune) and Anti-SARS (human convalescent sera) analysis of pGAP-1190 purified from *Pichia pastoris* supernatant (pre/post Endonuclease H treatment).

Figure 50 shows the results of mass spectroscopy (MALDI-ESI) of S glycoprotein expressed in *Pichia pastoris* (SEQ ID NO:33).

Figure 51A shows the results of SDS-PAGE and Coomassie blue staining of fractionated *Pichia pastoris*-derived rS glycoprotein (cA1) following diafiltration through a >300 kDa membrane cut-off. Ten µl of 10x concentrate was loaded. Figure 51B shows the immunoreactivity of clarified supernatant from a growing culture of cA1 material 48 hours following conversion from batch to fed-batch fermentation with two conformational dependent monoclonal antibodies.

Figure 52 shows the results of size exclusion HPLC over TSK SW4000$_{XL}$ (7.8 mm x 30 cm). The column was equilibrated with 0.1 M phosphate containing 0.25 M sodium chloride, pH 7.0 and appropriate size standards were included. Panel A shows a profile of diafiltered culture supernate harvested from cA1 fermentation. Fractionated samples were harvested and their immunoreactivity against the anti-SARS polyclonal (1:200) was evaluated in a dot blot format (5 µl/dot). Panel B shows the results of a re-folding study on soluble aggregate. Samples were normalized for HMW soluble aggregate.

Figure 53 shows determination of the molecular mass of fractionated fermentation samples by size exclusion HPLC over TSK SW4000$_{XL}$ coupled to a light scattering detector (Wyatt Technologies). The molar mass of selected peaks was calculated from the intensity of scattered light, times the square of the change in refractive index with respect to concentration. The separation range for this particular column is from 20,000 – 7,000,000 daltons.

Figure 54 shows Coomassie stain (SDS-PAGE; A) and Immunoblot (anti-SARS-CoV polyclonal; B) analysis of the expression of rS glycoprotein monomer in continuous culture.

Figure 55 shows native PAGE analysis of rS glycoprotein by Coomassie stain (PAGE; A) and Immunoblot (anti-SARS-CoV polyclonal hyperimmune).

Figure 56 is a graph showing SE-HPLC analysis of rS glycoprotein HMW complexes.

Figure 57 shows native PAGE and immunoreactivity profiling with SARS-specific antibodies.

Figure 58 is a graph showing the fractionation and immunoreactivity profile of HMW rS glycoprotein.

Figure 59 is a schematic representation of the vaccinia insertion vector pTK53-gpt-Spike. Abbreviations: Spike – SARS Spike gene; *gpt* – dominant selectable marker *E. coli* guanine phosphoribosiltransferase; P11, P7.5 – Vaccinia virus promoters; pUC – plasmid replication origin; tk_L and tk_R - left and right shoulders of thymidine kinase (tk) gene; EcoRI and BamHI – restriction endonuclease cleavage sites used for cloning.

Figure 60 is a schematic outline of the TDS approach used for generating rMVA-spike virus.

Figure 61 is a schematic outline of rMVA-spike studies.

Figure 62 shows Western blot analysis of rMVA-S (A, B, C, and D) and CEF/rMVA-N (1, 2, 3, and 4) cell lysates. MVA was grown in Chick Embryo Fibroblasts (CEF). The control is MVA-infected CEF.

Figure 63 provides a linear map of pTK53-N, as well as the nucleotide (SEQ ID NO:34) and amino acid (SEQ ID NO:35) sequences of the SARS nucleocapsid protein.

Figure 64 provides the nucleotide (SEQ ID NO:36) and amino acid (SEQ ID NO:37) sequence of a SARS spike protein.

Figure 65 provides the nucleotide sequence of a SARS coronavirus genome (SEQ ID NO:38).

Detailed Description

The invention relates to vaccines and methods that can be used to prevent or to treat Severe Acute Respiratory Syndrome (SARS) caused by human coronaviruses. Viruses causing this disease are known as human coronavirus/SARS, CoV-SARS, TOR2, and Urbani SARS-associated coronavirus. Also included in the invention are methods of producing proteins (e.g., spike proteins and nucleocapsid proteins) of human

coronaviruses causing SARS. as well as SARS spike and nucleocapsid proteins, and nucleic acid molecules encoding these proteins.

The vaccines of the invention can be used in methods to prevent SARS in patients, such as human patients. In these methods, one or more immunogenic agents derived from a human coronavirus causing SARS are administered to a patient. The agent(s) used can include, for example, an inactivated preparation of the virus or a fraction thereof, or an attenuated version of the virus. The agent(s) can also include an isolated protein (or fragment) from the virus or a nucleic acid molecule encoding such a protein. As a specific example, which is discussed in further detail below, the spike protein of a human coronavirus that causes SARS (or a nucleic acid molecule encoding such a protein) can be used in the vaccines of the invention. Also, the SARS nucleocapsid protein (or a nucleic acid molecule encoding such a protein) can be used. Further, these proteins or nucleic acid molecules (or immunogenic fragments thereof) can be used individually or together, optionally in combination with other agents, such as adjuvants.

The vaccines can also be used to treat patients that have already been exposed to or infected by a virus causing SARS. Optionally, such therapeutic vaccination can be carried out in conjunction with antiviral therapy involving, for example, administration of antiviral agents, such as oseltamivir or ribavirin. The therapeutic vaccines can also be administered with steroids, in combination with ribavirin and other antimicrobial agents.

As is noted above, spike proteins from human coronaviruses causing SARS can be used in the vaccines of the present invention. The nucleotide and amino acid sequences of one example of such a protein are provided herein as SEQ ID NOs:36 and 37, respectively (also see Figure 64). In addition, SARS nucleocapsid proteins can be used, and the nucleotide and amino acid sequences of an example of such a protein are provided in Figure 63 (SEQ ID NO:34 and SEQ ID NO:35). These sequences and fragments and variants thereof (see above) are also included in the invention. These sequences were identified in a sequence of an entire genome of a human coronavirus causing SARS (SEQ ID NO:38).

The proteins of the invention can be made, for example, using a eukaryotic or prokaryotic recombinant expression system. Eukaryotic hosts include, for example, yeast cells (e.g., *Pichia Pastoris* or *Saccharomyces cerevisiae*), mammalian cells (e.g., COS1, NIH3T3, HeLa, or JEG3 cells), arthropods cells (e.g., *Spodoptera frugiperda* (SF9) cells), and plant cells, while an example of a prokaryotic host is *E. coli*. Eukaryotic and prokaryotic cells for use in the invention are available from a number of different sources that are known to those skilled in the art, e.g., the American Type Culture Collection (ATCC; Manassas, Virginia; see also Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, 1998, which is hereby incorporated by reference). The method of transformation and the choice of expression vehicle (e.g., expression vector) will depend on the host system selected. Transformation and transfection methods, as well as expression vehicles, are described, e.g., in Ausubel et al., supra; also see, e.g., Pouwels et al., Cloning Vectors: A Laboratory Manual, 1985, Supp. 1987. Specific examples of expression systems that can be used in the invention are described further as follows.

Preferred expression systems for use in making the antigens of the invention are those in which post-translational glycosylation takes place, and include, for example, yeast, mammalian, and insect systems. This is particularly important with respect to SARS spike proteins, which are glycosylated (see below). Examples of yeast hosts that can be used in the invention include *Pichia pastoris*, *Pichia methanolica*, *Hansunela polymorpha*, *Schizosaccharomyces pombe*, and *Saccharomyces cerevisiae*. In the case of *P. pastoris*, specific examples of host strains that can be used include X-33, GS115, KM71, KM71H, SMD1168, and SMD1168H. Examples of yeast vectors that can be used include pPIC vectors (Invitrogen), such as pPICZalpha for secretion using the alpha factor secretion signal. Also, pPIC vectors that allow multi-copy integrants can be used. These vectors allow multiple insertions into the genome. Use of methalymine or methanol-inducible expression systems can also be used. In another example of a yeast-based system that can be used in the invention, the yeast used to produce the proteins are engineered to make proteins so that they are glycosylated similarly to human proteins (see, e.g., Hamilton et al., Science 301:1244-1246, 2003).

Transient transfection of a eukaryotic expression plasmid containing a spike or nucleocapsid protein gene into a mammalian host cell (e.g., COS1, NIH3T3, HeLa, or JEG3 cells) allows the transient production of the protein by the transfected host cell. The proteins can also be produced by a stably-transfected eukaryotic (e.g., mammalian) cell line. A number of vectors suitable for stable transfection of mammalian cells are available to the public (see, e.g., Pouwels et al., supra), as are methods for constructing lines including such cells (see, e.g., Ausubel et al., supra). In one example, cDNA encoding a spike or nucleocapsid protein, fusion, mutant, or polypeptide fragment is cloned into an expression vector that includes the dihydrofolate reductase (DHFR) gene. Integration of the plasmid and, therefore, integration of the protein-encoding gene, into the host cell chromosome is selected for by inclusion of 0.01-300 μM methotrexate in the cell culture medium (Ausubel et al., supra). This dominant selection can be accomplished in most cell types. Recombinant protein expression can be increased by DHFR-mediated amplification of the transfected gene. Methods for selecting cell lines bearing gene amplifications are described in Ausubel et al., supra. These methods generally involve extended culture in medium containing gradually increasing levels of methotrexate. The most commonly used DHFR-containing expression vectors are pCVSEII-DHFR and pAdD26SV(A) (described, for example, in Ausubel et al., supra). The host cells described above or, preferably, a DHFR-deficient CHO cell line (e.g., CHO DHFR- cells, ATCC Accession No. CRL 9096) are among those that are most preferred for DHFR selection of a stably transfected cell line or DHFR-mediated gene amplification.

Another preferred eukaryotic expression system is the baculovirus system using, for example, the vector pBacPAK9, which is available from Clontech (Palo Alto, CA). If desired, this system can be used in conjunction with other protein expression techniques, for example, the myc tag approach described by Evan et al. (Molecular and Cellular Biology 5:3610-3616, 1985). Additional examples of insect systems that can be used are the Bac-to-Bac Baculovirus expression system, employing, e.g., pFastBac1 vectors, as well as a Drosophila expression system employing S2 cells (see below). The latter

system can employ, for example, the pMT/Bip/V5-His vector for regulated, secreted expression.

Expression of foreign molecules in bacteria, such as *Escherichia coli*, requires the insertion of a foreign nucleic acid molecule, e.g., a spike nucleic acid molecule or a nucleocapsid nucleic acid molecule, into a bacterial expression vector. Such plasmid vectors include several elements required for the propagation of the plasmid in bacteria, and for expression of foreign DNA contained within the plasmid. Propagation of only plasmid-bearing bacteria is achieved by introducing, into the plasmid, a selectable marker-encoding gene that allows plasmid-bearing bacteria to grow in the presence of an otherwise toxic drug. The plasmid also contains a transcriptional promoter capable of directing synthesis of large amounts of mRNA from the foreign DNA. Such promoters can be, but are not necessarily, inducible promoters that initiate transcription upon induction by culture under appropriate conditions (e.g., in the presence of a drug that activates the promoter). The plasmid also, preferably, contains a polylinker to simplify insertion of the gene in the correct orientation within the vector. An example of a prokaryotic system that can be used is *E. coli*, using BL21 lambda DE3 and pET vectors, pET26 with a pelB leader for expression to the periplasm, or pET24 for expression of native protein or overlapping fragments thereof.

Proteins of the invention can also be obtained using *in vitro* methods. For example, *in vitro* expression of the proteins, fusions, polypeptide fragments, or mutants encoded by cloned DNA can also be carried out using the T7 late-promoter expression system. This system depends on the regulated expression of T7 RNA polymerase, an enzyme encoded in the DNA of bacteriophage T7. The T7 RNA polymerase initiates transcription at a specific 23 base pair promoter sequence called the T7 late promoter. Copies of the T7 late promoter are located at several sites on the T7 genome, but none are present in *E. coli* chromosomal DNA. As a result, in T7-infected *E. coli*, T7 RNA polymerase catalyzes transcription of viral genes, but not *E. coli* genes. In this expression system, recombinant *E. coli* cells are first engineered to carry the gene encoding T7 RNA polymerase next to the *lac* promoter. In the presence of IPTG, these cells transcribe the T7 polymerase gene at a high rate and synthesize abundant amounts

of T7 RNA polymerase. These cells are then transformed with plasmid vectors that carry a copy of the T7 late promoter protein. When IPTG is added to the culture medium containing these transformed *E. coli* cells, large amounts of T7 RNA polymerase are produced. The polymerase then binds to the T7 late promoter on the plasmid expression vectors, catalyzing transcription of the inserted cDNA at a high rate. Since each *E. coli* cell contains many copies of the expression vector, large amounts of mRNA corresponding to the cloned cDNA can be produced in this system and the resulting protein can be radioactively labeled.

Plasmid vectors containing late promoters and the corresponding RNA polymerases from related bacteriophages, such as T3, T5, and SP6, can also be used for *in vitro* production of proteins from cloned DNA. *E. coli* can also be used for expression using an M13 phage, such as mGPI-2. Furthermore, vectors that contain phage lambda regulatory sequences, or vectors that direct the expression of fusion proteins, for example, a maltose-binding protein fusion protein or a glutathione-S-transferase fusion protein, also can be used for expression in *E. coli*.

Polypeptides of the invention, particularly short fragments and longer fragments of the N-terminus and C-terminus of the proteins, can also be produced by chemical synthesis (e.g., by the methods described in Solid Phase Peptide Synthesis, 2nd ed., 1984, The Pierce Chemical Co., Rockford, IL). These general techniques of polypeptide expression and purification can also be used to produce and isolate useful fragments or analogs, as described herein.

Once an appropriate expression vector containing a gene, or a fragment, fusion, or mutant thereof, is constructed, it can be introduced into an appropriate host cell using a transformation technique, such as, for example, calcium phosphate transfection, DEAE-dextran transfection, electroporation, microinjection, protoplast fusion, or liposome-mediated transfection. Host cells that can be transfected with the vectors of the invention can include, but are not limited to, *E. coli* or other bacteria, yeast, fungi, insect cells (using, for example, baculoviral vectors for expression), or cells derived from mice, humans, or other animals (see, e.g., above). Mammalian cells can also be used to express the proteins of the invention using a virus expression system (e.g., a vaccinia virus

expression system) described, for example, in Ausubel et al., supra. As a specific example of a vaccinia virus system that can be used, see, e.g., Moore et al., EMBO J., 11:1973-1980, 1992, erratum at EMBO J. 11:3490, 1992; Skinner et al., J. Gen. Virol. 75:2495-2498, 1994; and Sroller et al., Arch. Virol. 143:1311-1320, 1998, which describe the use of a Modified Vaccinia Ankara (MVA) strain. Also see, e.g., U.S. Patent No. 6,440,422.

Upon expression, a recombinant polypeptide of the invention (or a polypeptide derivative) is produced and remains in the intracellular compartment, is secreted/excreted in the extracellular medium or in the periplasmic space, or is embedded in the cellular membrane. Preferably, the polypeptide is secreted. The polypeptide can then be recovered in a substantially purified form from the cell extract or from the supernatant after centrifugation of the cell culture. Typically, the recombinant polypeptide can be purified by antibody-based affinity purification or by any other method known to a person skilled in the art, such as by genetic fusion to a small affinity-binding domain. Antibody-based affinity purification methods are also available for purifying a polypeptide of the invention. Antibodies useful for immunoaffinity purification of the polypeptides of the invention can be obtained using standard.

As is discussed further below, we have found that certain spike proteins produced using the methods described herein assemble into trimeric structures, which have been observed to form with certain spike proteins from animal coronaviruses. Thus, the invention includes human coronavirus spike proteins in this form, as well as monomeric and dimeric forms, and the use of the proteins in such forms in the methods described herein.

In addition to protein based antigens, the methods of the invention can employ nucleic acid (e.g., DNA or RNA)-based antigens, whether in the form of a vector delivering a gene to be expressed or administration of a nucleic acid molecule itself. Polynucleotides of the invention can also be used in DNA vaccination methods, using either a viral or bacterial host as gene delivery vehicle (live vaccine vector) or administering the gene in a free form, e.g., inserted into a plasmid. Typically, a DNA molecule is placed under the control of a promoter suitable for expression in a

mammalian cell. The promoter can function ubiquitously or tissue-specifically. Examples of non-tissue specific promoters include the early Cytomegalovirus (CMV) promoter (U.S. Patent No. 4,168,062) and the Rous Sarcoma Virus promoter (Norton et al., Molec. Cell Biol. 5:281, 1985). The desmin promoter (Li et al., Gene 78:243, 1989; Li et al., J. Biol. Chem. 266:6562, 1991; Li et al., J. Biol. Chem. 268:10403, 1993) is tissue-specific and drives expression in muscle cells. More generally, useful promoters and vectors are described, e.g., in WO 94/21797 and by Hartikka et al. (Human Gene Therapy 7:1205, 1996).

Live vaccine vectors that can be used in the invention include viral vectors, such as adenoviruses and poxviruses (e.g., vaccinia virus vectors, such as MVA vectors), as well as bacterial vectors, e.g., Shigella, Salmonella, Vibrio cholerae, Lactobacillus, Bacille bilié de Calmette-Guérin (BCG), and Streptococcus. An example of an adenovirus vector, as well as a method for constructing an adenovirus vector capable of expressing a polynucleotide molecule of the invention, is described in U.S. Patent No. 4,920,209. Poxvirus vectors that can be used in the invention include, e.g., vaccinia and canary pox viruses, which are described in U.S. Patent No. 4,722,848 and U.S. Patent No. 5,364,773, respectively (also see, e.g., Tartaglia et al., Virology 188:217, 1992, for a description of a vaccinia virus vector, and Taylor et al, Vaccine 13:539, 1995, for a description of a canary poxvirus vector). Poxvirus vectors capable of expressing a polynucleotide of the invention can be obtained by homologous recombination, as described in Kieny et al. (Nature 312:163, 1984) so that the polynucleotide of the invention is inserted in the viral genome under appropriate conditions for expression in mammalian cells. Details of the use of a pox-based vector are provided in the Examples, below.

In addition to viral-based vectors, bacterial vectors can be used in the invention to administer SARS proteins. Attenuated *Salmonella typhimurium* strains, genetically engineered for recombinant expression of heterologous antigens, and their use as oral vaccines, are described by Nakayama et al. (Bio/Technology 6:693, 1988) and in WO 92/11361. Preferred routes of administration for these vectors include all mucosal routes (e.g., intranasal or oral routes). Others bacterial strains useful as vaccine vectors

are described by High et al. (EMBO 11:1991, 1992) and Sizemore et al. (Science 270:299, 1995; *Shigella flexneri*); Medaglini et al. (Proc. Natl. Acad. Sci. U.S.A. 92:6868, 1995; (*Streptococcus gordonii*); Flynn (Cell. Mol. Biol. 40 (suppl. I):31, 1194), and in WO 88/6626, WO 90/0594, WO 91/13157, WO 92/1796, and WO 92/21376 (Bacille Calmette Guerin). In bacterial vectors, a polynucleotide of the invention can be inserted into the bacterial genome or it can remain in a free state, for example, carried on a plasmid. An adjuvant can also be added to a composition containing a bacterial vector vaccine.

Methods for administering vaccine compositions including the proteins, fragment, nucleic acid molecules, or vectors of the invention are described as follows.

Administration

As is noted above, the vaccines of the invention can include SARS spike or nucleocapsid polypeptides or immunogenic fragments, or nucleic acid molecules encoding such polypeptides or immunogenic fragments. The vaccines can be administered using routes, regimens, and formulations determined to be appropriate by those of skill in this art. Examples of these and other parameters for consideration in administering the vaccines of the invention are discussed as follows.

The vaccines of the invention can be administered by any conventional route in use in the vaccine field, for example, by a parenteral (e.g., subcutaneous, intradermal, intraepidermal, intramuscular, intravenous, or intraperitoneal) or a mucosal (e.g., ocular, intranasal, oral, gastric, pulmonary, intestinal, rectal, vaginal, or urinary tract) route.

Appropriate amounts of vaccine to be administered can readily be determined by those of skill in the art, and can depend upon various parameters such as the nature of the vaccine vector itself, the route and frequency of administration, the presence/absence of adjuvant, the desired effect (e.g., protection and/or treatment), and the condition of the mammal to be vaccinated (e.g., the weight, age, and general health of the mammal). In general, 0.1 μg - 1 mg, e.g., 1-500 μg, e.g., or 10-100 μg (e.g., 20-80, 30-70, 40-60 or about 50 μg), can be administered. For example, a vaccine of the invention can be administered mucosally in an amount ranging from about 10 μg to about 500 mg, e.g.,

from about 1 mg to about 200 mg. For a parenteral route of administration, the dose usually should not exceed about 1 mg, and can be, preferably, about 50-500, e.g., 100-250 μg.

The vaccines of the invention can be administered in regimens that can be determined to be appropriate by those of skill in this art. For example, the administration can be achieved in a single dose or repeated at intervals. As a specific example, the vaccines can be administered in three doses biweekly, 1 month apart, or on days 0, 28, and 56 of a multi-dose regimen. In another example, a priming dose is followed by 1-3 booster doses at weekly or monthly intervals (e.g., a boost within 1-6 months), with follow-up boosting every 1-5 (e.g., 3) years, if needed. As yet another example, a subject can initially be primed with a vaccine vector of the invention, such as a pox virus (e.g., MVA or adenovirus) by, e.g., a parenteral route, and then boosted (e.g., 2-4 times) with a polypeptide encoded by the vaccine vector by the parenteral or mucosal route. Alternatively, a polypeptide can be used in a priming step, and boosting can be carried out using a vaccine vector, such as a pox virus or an adenovirus. In another example, liposomes associated with a polypeptide or polypeptide of the invention can be used for priming, with boosting being carried out mucosally using a soluble polypeptide or polypeptide derivative of the invention, in combination with a mucosal adjuvant (e.g., LT). In a further example, the antigen is administered mucosally (e.g., intranasally) in a priming step, and boosting is by parenteral administration. Further, the vaccines described herein can be used in combination with each other or other vaccines against SARS, by co-administration or in prime/boost methods in which a vaccine as described herein is used in either the prime or a boosting step, and the other vaccine is used in a step in which a vaccine as described herein is not used.

The vaccines of the invention can be formulated using standard methods (see, e.g., in *Remington's Pharmaceutical Sciences* (18^{th} edition), ed. A. Gennaro, 1990, Mack Publishing Company, Easton, PA). In addition to the antigenic agent(s), the vaccines can optionally also include an adjuvant. Examples of adjuvants that can be included in the vaccines of the invention include alum and other aluminum compounds (e.g., aluminum hydroxide, aluminum phosphate, and aluminum hydroxy phosphate), DC-Chol, QS-21,

MPL, Ribi, as well as other parenteral adjuvants that are known in the art. Additional formulations that can be used can include the use of liposomes, such as neutral or anionic liposomes, microspheres, or virus-like particles (VLPs), to facilitate delivery and/or enhance the immune response. Another example of an adjuvant that can be used is ISCOMs, which can be used, e.g., in mucosal (e.g., intranasal or oral) administration of, e.g., the polypeptide antigens described herein. These compounds are readily available to those skilled in the art; for example, see Liposomes: A Practical Approach (supra).

Additional adjuvants that can be used for mucosal administration include, for example, bacterial toxins, e.g., the cholera toxin (CT), the *E. coli* heat-labile toxin (LT), the *Clostridium difficile* toxin A, the pertussis toxin (PT), and combinations, subunits, toxoids, or mutants thereof. For example, a purified preparation of native cholera toxin subunit B (CTB) can be used. Fragments, homologs, derivatives, and fusions to any of these toxins can also be used, provided that they retain adjuvant activity. Preferably, a mutant having reduced toxicity is used. Suitable mutants are described, e.g., in WO 95/17211 (Arg-7-Lys CT mutant), WO 96/6627 (Arg-192-Gly LT mutant), and WO 95/34323 (Arg-9-Lys and Glu-129-Gly PT mutant). Additional LT mutants that can be used in the methods and compositions of the invention include, e.g., Ser-63-Lys, Ala-69-Gly, Glu-110-Asp, and Glu-112-Asp mutants. Other adjuvants, such as the bacterial monophosphoryl lipid A (MPLA) of, e.g., *E. coli*, *Salmonella minnesota*, *Salmonella typhimurium*, or *Shigella flexneri*; saponins, and polylactide glycolide (PLGA) microspheres, can also be used in mucosal administration. Adjuvants useful for both mucosal and parenteral administrations, such as polyphosphazene (WO 95/2415), can also be used.

As is noted above, the vaccination methods of the invention can also include the use of polynucleotide molecules, which can, optionally, be administered in a vector. A polynucleotide of the invention can be used in a naked form, free of any delivery vehicles, such as anionic liposomes, cationic lipids, microparticles, e.g., gold microparticles, precipitating agents, e.g., calcium phosphate, or any other transfection-facilitating agent. In this case, the polynucleotide can simply be diluted in a physiologically acceptable solution, such as sterile saline or sterile buffered saline, with

or without a carrier. When present, the carrier preferably is isotonic, hypotonic, or weakly hypertonic, and has a relatively low ionic strength, such as provided by a sucrose solution, e.g., a solution containing 20% sucrose.

Alternatively, a polynucleotide can be associated with agents that assist in cellular uptake. It can be, e.g., (i) complemented with a chemical agent that modifies cellular permeability, such as bupivacaine (see, e.g., WO 94/16737), (ii) encapsulated into liposomes, or (iii) associated with cationic lipids or silica, gold, or tungsten microparticles. Anionic and neutral liposomes are well-known in the art (see, e.g., Liposomes: A Practical Approach, RPC New Ed, IRL Press, 1990, for a detailed description of methods for making liposomes) and are useful for delivering a large range of products, including polynucleotides.

Cationic lipids can also be used for gene delivery. Such lipids include, for example, Lipofectin™, which is also known as DOTMA (N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride), DOTAP (1,2-bis(oleyloxy)-3-(trimethylammonio)propane), DDAB (dimethyldioctadecylammonium bromide), DOGS (dioctadecylamidologlycyl spermine), and cholesterol derivatives. A description of these cationic lipids can be found in EP 187,702, WO 90/11092, U.S. Patent No. 5,283,185, WO 91/15501, WO 95/26356, and U.S. Patent No. 5,527,928. Cationic lipids for gene delivery are preferably used in association with a neutral lipid such as DOPE (dioleyl phosphatidylethanolamine; WO 90/11092). Other transfection-facilitating compounds can be added to a formulation containing cationic liposomes. A number of them are described in, e.g., WO 93/18759, WO 93/19768, WO 94/25608, and WO 95/2397. They include, e.g., spermine derivatives useful for facilitating the transport of DNA through the nuclear membrane (see, for example, WO 93/18759) and membrane-permeabilizing compounds such as GALA, Gramicidine S, and cationic bile salts (see, for example, WO 93/19768).

Gold or tungsten microparticles can also be used for gene delivery, as described in WO 91/359, WO 93/17706, and by Tang et al. (Nature 356:152, 1992). In this case, the microparticle-coated polynucleotides can be injected via intradermal or

intraepidermal routes using a needleless injection device ("gene gun"), such as those described in U.S. Patent No. 4,945,050, U.S. Patent No. 5,015,580, and WO 94/24263.

The amount of DNA to be used in a vaccine recipient depends, e.g., on the strength of the promoter used in the DNA construct, the immunogenicity of the expressed gene product, the condition of the mammal intended for administration (e.g., the weight, age, and general health of the mammal), the mode of administration, and the type of formulation. In general, a therapeutically or prophylactically effective dose from about 1 μg to about 1 mg, preferably, from about 10 μg to about 800 μg, and, more preferably, from about 25 μg to about 250 μg, can be administered to human adults. The administration can be achieved in a single dose or repeated at intervals.

A preferred approach for vaccination according to the present invention involves the use of a live vector, such as a live viral vector. For example, nucleotide sequences encoding SARS spike proteins or immunogenic fragments thereof, as described elsewhere herein, can be inserted into a live vector, such as a pox vector, which is administered in vaccination methods. Additional viral and bacterial vectors that can be used in the invention are known in the art (also, see above). As a specific example, the attenuated vaccinia virus Modified Vaccinia Ankara (MVA) can be used as a viral delivery vehicle in the invention. Details of the use of such a viral vector are provided below in the Examples. In general, the dose of a viral vector vaccine, for therapeutic or prophylactic use, can be from about $1x10^4$ to about $1x10^{11}$, e.g., $1x10^7$ to $1x10^{10}$, or $1x10^7$ to about $1x10^9$, plaque-forming units per kilogram. Such vectors can be administered, e.g., parenterally, for example, in 3 doses that are 4 weeks apart.

Also included in the invention are passive immunization methods for preventing or treating SARS infection. In these methods, antibodies against the SARS virus, or one or more components thereof, are administered to patients to prevent or treat infection. As a specific example, polyclonal hyperimmune globulin that is obtained from plasma donors that have been actively immunized with a SARS antigen (e.g., a spike protein antigen or a nucleocapsid protein antigen; see, e.g., above) can be used. Routes of administration include, for example, mucosal and parenteral routes. For example, in the case of mucosal administration, the antibody preparation can be administered in the

form of nose drops or by inhalation, using standard methods in the art. Other mucosal routes, such as those listed above, can also be used. In the case of parenteral administration, subcutaneous injection or any other parenteral route (see, e.g., those listed above) can be used. The passive immunization methods can be used as sole approaches to prevention or treatment, or can be used in combination with active vaccination approaches, such as those described herein (see, e.g., WO 99/20304 for additional details on passive immunization approaches).

Examples

Example 1 - Expression Constructs

Constructs for expressing the SARS coronavirus spike proteins in three different eukaryotic systems, the yeast *Pichia pastoris*, mammalian CHO cells, and drosophila S2 cells, were made and characterized. The details of these constructs are summarized in the following table, and are illustrated in Figures 1-38. The constructs each lack the native N-terminal spike signal sequence (amino acids 1-13), in favor of those provided by the vectors used in each of the systems. The vector-provided signal sequences ensure that the proteins are secreted in the relevant systems.

As is illustrated in Figure 39, the SARS spike protein can be divided up into an extracellular domain, a transmembrane domain, and a cytoplasmic tail. We tested constructs that include different combinations of these regions. For example, the 14-719 (or 709) constructs include the extracellular domain (i.e., the putative S1 domain, which represents the receptor binding domain and the region including neutralization determinants); the 14-883 constructs include the extracellular domain and the S2 domain, but not the intracellular coiled coil domain, while the 14-1190 constructs include the extracellular domain, but not the transmembrane domain, and the cytoplasmic tail. The table set forth below provides information as to the vectors used, the construct names, and the spike protein amino acids included in the constructs, for each of the three systems. Constructs including amino acids 14-1190, 14-883, and 14-709 are alternatively referred to herein as clones A1, A2, and A3, respectively.

***Pichia pastoris* expression constructs**

Expression system	Vector	Construct name	Spike protein limits (aa)
Inducible	pPICZalpha	P1-2	14-709
	pPICZalpha	P1-2.2	14-719
	pPICZalpha	P3-10 mutant (H641R)	14-883
	pPICZalpha	P3-10.2	14-883
	pPICZalpha	P5-12*	14-1190
* contains 2 silent mutations within aa C133, C1108			
Constitutive	pGAPZalpha	G1-8	14-709
	pGAPZalpha	G1-8.2	14-719
	pGAPZalpha	G3-7 mutant (Y484C)	14-883
	pPAGZalpha	G3-7.2	14-883
	pPAGZalpha	G5-14**	14-1190
** contains 2 silent mutations within aa Y723, G739			

pSEC expression constructs (CHO cells)

Expression system	Vector	Construct name	Spike protein limits (aa)
Constitutive	pSEC	pSEC-719	14-719
	pSEC	pSEC-883	14-883
	pSEC	pSEC-1190	14-1190

DES constructs (Drosophila S2 cells)

Expression system	Vector	Construct name	Spike protein limits (aa)
Inducible	pMT	pMT-719	14-719
	pMT	pMT-883	14-883
	pMT	pMT-1190	14-1190

Additional details as to the construction of these constructs are as follows. Crude RNA was extracted from Vero E6 cells infected with SARS-CoV (provided by the CDC) and reverse transcriptase-polymerase chain reaction (RT-PCR) was performed for cloning. cDNA clones representing all structural genes in their entirety were constructed and characterized by DNA sequencing. Clones A1-A3 were constructed by PCR in the pPICZalpha and pGAPZalpha expression vectors for inducible and constitutive expression in *Pichia pastoris*, respectively. Fragments were engineered in-frame at the N-terminus with the alpha factor signal sequence allowing for export from a growing culture and were prematurely terminated at the C-terminus with a stop codon. Clones A1-A3 were then electroporated into competent X-33 *Pichia pastoris* and transformants were evaluated for integration by PCR, copy number by enhanced resistance to the selectable marker carried on the integration vector, and expression by, e.g., immunoreactivity with SARS-specific antisera using a dot blot format (see below).

Example 2 - Expression Studies

The constructs described above were analyzed for expression in the relevant systems, with the goal being to analyze the systems for yield, purity, solubility, and glycosylation. After such initial characterization, virus neutralization studies in, e.g., mice, can be carried out to determine appropriate regimens, doses, scheduling, adjuvants, and formulations, and then efficacy can be confirmed, if desired, in an appropriate non-human primate model. In each system, clones were generated by introduction of the constructs noted above into cells by lipofection or electroporation (Figure 40).

A generalized strategy for constitutive (CHO) and inducible (S2) expression of recombinant spike proteins is illustrated in Figure 41. Briefly, the spike gene is cloned into an appropriate vector (e.g., pMT/BiP for S2 cells or pSec/FRT TOPO for Flp-In CHO cells), positive transformants are selected and sequenced, and then the constructs are integrated into the S2 or CHO cells by use of a co-transfected recombination plasmid and selection with hygromycin (CHO) or Blasticidine (S2). The integrants are then screened for high level expression, a candidate is selected, expression is optimized, and production is then scaled up, if desired. Figure 42 shows the results of PCR screening of genomic DNA purified from transiently transfected S2 cells 24 and 48 hours after transfection with pMT-719, pMT-883, and pMT-1190 constructs, as well as Western blot analysis of these cells. Figure 43 presents RT-PCR data showing that the spike 719, 883, and 1190 genes are expressed in CHO cells.

The generalized strategy for expression of recombinant spike proteins in the yeast *Pichia pastoris* is illustrated in Figure 44. Briefly, constructs are sequenced, midi-prepped, and subcloned, and then are integrated into *P. pastoris* by linearization, electroporation, and Zeocin selection. The integrants are then screened for high copy numbers, fermented, and a candidate is selected and optimized. Figure 45 shows spike gene-specific PCR of chromosomal DNA, confirming integration for constructs encoding 1190 and 883 amino acids of the SARS spike protein, as described above. In particular, Figure 45 shows a sample set of PCR positive (N-terminal fragment) integrants for A1 (1190) and A2 (883) constructs for both inducible and constitutive expression. Small-scale expression studies were then performed on integrants to identify clones for bench-

scale fermentation. Figure 45 further shows the immunoreactivity of a panel of A1 (1190) integrants engineered to produce full-length ectodomain following inducible expression, that are immunoreactive with a neutralizing, murine hyperimmune polyclonal antibody raised against gamma-irradiated SARS-CoV. Clone 64 (identified by the arrow) was observed to react strongly with the SARS polyclonal serum and was selected for further study. Similar studies identified a clone that expressed immunoreactive product following constitutive expression.

Figure 46 shows the results of analysis of clones constitutively expressing the 1190 (lanes 2 and 6) and 883 (lanes 3 and 7) amino acid versions of recombinant spike. Panel A is a glycostain; panel B is an immunoblot with an anti-SARS coronavirus murine polyclonal antibody; panel C is a dot blot using such an antibody; panel D is an immunoblot using human convalescent sera; and panel E is a dot blot using the latter sera. In lanes 2 and 3, the proteins are glycosylated, while in lanes 6 and 7, the proteins have been treated with Endonuclease H, resulting in deglycosylation. These data show that high molecular weight material is immunoreactive with anti-spike antibodies, and that this material breaks down upon Endonuclease H treatment, yielding the expected 139 kDa (full ectodomain, 1190) and 98 kDa (bulbar head, 883) products. The dot blot results show the maintenance of at least some conformational integrity of the recombinant proteins.

The spike protein was purified from fermentor bulk material by successive diafiltration steps (>300 kDa, 15x; 100-300 kDa, 15x; and <100 kDa), and the fractions were tested for immunoreactivity with an anti-SARS coronavirus antibody (Figure 47). Most of the immunoreactivity was found in the >300 kDa fraction. No reactivity was observed with lower molecular weight material that could represent monomer or similarly treated material expressed from the control strain X-33 lacking the S gene.

The partially purified retentate material (1190) was then purified further by lectin affinity chromatography in batch mode by binding to Concanavalin A-Sepharose 4B and eluting with sugar (methyl a-D-Mannopyranoside, 750 mM) (Figure 48). The eluted material (glycosylated and deglycosylated samples) was fractionated by SDS-PAGE, and detected by Western blot analysis as high molecular weight material at about 180-250

kDa (glycosylated), while Endo-H treated material (deglycosylated) was detected at about 138 kDa, as was expected. These studies show that the spike protein secreted from *P. pastoris* through the secretory pathway was glycosylated.

The results of additional studies showing the purification of pGAP-1190 from *P. pastoris* supernatants are shown in Figure 49. SDS-PAGE and removal of DTT from the sample buffer suggests the presence of monomeric, pichia glycosylated spike protein of >250 kDa. Figure 50 shows that mass spectrometry (MALDI-ESI) confirms expression of the S glycoprotein in pichia.

In other studies, fed-batch fermentation (2 L) of *Pichia pastoris* integrants expressing full-length rS ectodomain (cA1 and iA1) was performed in a controlled environment with basal salts medium (BSM) in the absence of selection. The following illustrates the methodology employed when performing constitutive expression of cA1 expressing full-length rS glycoprotein. Briefly a vial of cA1 research cell bank (RCB) was seeded into 100 ml of BSM plus PTM_1 trace salts and grown for 18 hours at 28°C. On day 2, a 5% inoculum was added to the fermentor containing BSM (pH5) containing 4% glycerol and grown at 30°C, with dissolved O_2 (DO) active control maintained at 35%, agitation set 100-1000 rpm, with airflow at 3.0 L/minute. Fermentation was monitored using BioCommand software (New Brunswick). At carbon exhaustion, the feed program was initiated were 50% glycerol solution was added with PTM_1 trace salts at a rate of 0.5%/liter/hour. After 2 additional days of fermentation, the culture supernate was harvested and EDTA added to 5 mM. Using fed-batch fermentation, we observed dramatic increases in production of target protein (Figure 51) with yields of monomer calculated at 100 mg/L by densitometry. Fermentation at low pH is known to be optimal for yeast growth and likely limited proteolytic breakdown of expressed product. We have achieved cell densities (OD_{600}) of > 400 using both our inducible and constitutive expression systems. Gel extraction of the 180 kDa monomer was confirmed as Spike protein following Mass Spectroscopy and in gel digestion and sequencing of generated fragments. Over 90% of the harvested sequence spanned the entire length of the protein confirming high-level expression of monomeric S glycoprotein.

Size exclusion. high-pressure liquid chromatography (HPLC) of diafiltered (> 300 kDa) material supported expression of several HMW species ranging in size from 100 to >1000 kDa (Figure 52A). Isolated fractions representing peaks i & ii and that ran near the void volume were demonstrated to be immunoreactive with the polyclonal anti-SARS hyperimmune raised against SARS-CoV (Figure 52A). Fractions 13-18 were subsequently treated with endo-deglycosidase H and in every case immunoreactivity with the mouse polyclonal was observed with a circa 130 kDa deglycosylated protein, as expected. Denaturation of the diafiltered material with Guanidine Hydrocholoride (GuHCl) plus DTT resolved much of the void volume peaks to the lower molecular weight species, peak iii (Figure 52B). When the denatured sample was dialyzed against citrate buffer (pH 4), the lower molecular weight component appeared to re-associate preferentially to the higher molecular weight peak ii (Figure 52B). This material was soluble and stable at +4°C. Dot blots of isolated fractions with both the mouse polyclonal and conformational dependent monoclonal antibodies, previously demonstrated to neutralize SARS-CoV, confirmed the immunoreactivity and structure of re-associated peak ii.

Different fractions representing peaks i, ii, and iii were then re-injected and further analyzed by light scattering with accurate molecular weight determinations of 160, 322, and 623 kDa, supporting monomeric, dimeric, and trimeric forms (Figure 53). Importantly, the trimeric structure was quite stable. The ability to correctly re-fold trimer as determined by immunoreactivity provides us with a method enabling enrichment of a preferred S glycoprotein structure.

Additional data supporting fermentation and expression of full-length rS ectodomain using continuous culture are described as follows. Constitutive expression of rS glycoprotein has been achieved for 40 days and the effect of temperature and pH on production of target protein monitored. The data supports expression of a soluble, rS glycoprotein at pH 7.0. To mitigate effects of enhanced proteolysis at the elevated pH due to intrinsic proteases produced by *Pichia pastoris,* fermentation is carried out at 15°C. Figure 54 shows expression of a circa 180 kDa monomer in neat culture supernatants and its immunoreactivity

profile with the murine polyclonal antibody raised against γ-irrdadiated SARS-CoV under denaturing/reducing conditions (Figures 54A and 54B). Two time-points are represented. Yields of rS glycoprotein using a constitutive expression system seem to favor increased production levels and is therefore the method of choice for production purposes. Performing continuous culture with this construct will ensure production of very respectable levels of rS glycoprotein for manufacturing purposes.

Previous data supported the expression of a high molecular weight (HMW) complex (> 300 kDa) that was preferentially immunoreactive with both anti-SARS-CoV polyclonal and monoclonal neutralizing antibodies. Polyacrylamide gel electrophoresis (PAGE) under native conditons supports the existence of a HMW complex and at least two isoforms of the rS glycoprotein. To better observe this phenomenon, culture supernatant was diafiltered through a 100 kDa membrane and concentrated 10-fold prior to running native PAGE (Figure 55, lanes 2). In the presence of reducing agents (DTT, lane 4; β-mercaptoethanol, lane 5) the higher of the two protein complexes is reduced to a single species, and presumably represents the monomeric form of the rS glycoprotein ectodomain. Both isomers are immunoreactive with the anti-SARS polyclonal antibody. The existence of the higher molecular weight isomer was also confirmed following re-folding of a gel-extracted 180 kDa monomer.

Following the concentration step over a 100 kDa membrane, recent studies have focused on separation of the HMW immunoreactive protein complexes by gel filtration to better define the immunoreactive products. Concentrated culture supernatant (Figure 55, lane 2) was first separated over Sephacryl S-500 (Pharmacia; Figure 55, lane 3) and both the void volume and one isolated peak further characterized by size exclusion high pressure liquid chromatography (SE-HPLC; Figure 56) over TSKSW 4000. Both samples were fractionated to discrete peaks, and harvested samples prepared for native PAGE, immunoreactivity profiling, and size determination by light scattering.

The initial S-500 gel filtration step successfully separated HMW protein complex from the lower molecular weight products (double headed arrow). Further separation of the low molecular weight proteins (B) using TSK SW4000 confirmed our ability to successfully fractionate the majority of the lower molecular weight products (fractions 12-19; Figure 57A) and enabled us to identify those protein complexes that were more immunoreactive (Figures 57B and 57C). A circa 300 kDa protein was isolated from fractions 12 and 13 that appeared to be preferentially recognized by both the polyclonal antibody raised against SARS-CoV and a monospecific polyclone (directed to a linear determinant on the C-terminus of the protein). Results are presented in a dot blot format and a Western blot of native PAGE.

Fractions 14 through 19 also appeared to be recognized with Spike protein-specific antibodies but to a lesser degree, possibly suggesting proteolytic cleavage to lower molecular weight products. Recent data supports this hypothesis through mass spectoscopy and peptide sequencing of gel extracted fragments representing these lower molecular weight peaks. Samples representing fractions 12 and 13 were then re-injected, SE-HPLC performed over TSK SW4000, and molecular weight determinations made by light scattering. A molecular weight of circa 300 kDa was assigned to the protein present in fraction 12 and probably represents a dimer. Fraction 13 was determined to have 2 proteins sizing at 300 and 177 kDa, likely representing the dimeric and monomeric forms, respectively.

Similar studies were performed on concentrated culture medium fractionated using Sephacryl S-300 to characterize the HMW complex. The first of two peaks harvested was further fractionated under pressure using TSK SW4000 (Figure 58). Samples were analyzed for their immunoreactivity against SARS-CoV and Spike protein specific antibodies and molecular weight determinations were made by light scattering. Immunoblot data confirmed the immunoreactivity of the dimer as determined in the LMW fractionation study, but also confirmed the existence of a third protein complex (peak 2) that was even

more immunoreactive with the anti-SARS-CoV antibody. Molecular weight determinations by light scattering supports peak 2 ranging in size from 450 – 750 kDa. Our previous studies support the existence of a trimer. Currently we believe the majority of protein in the medium to be Spike protein with total protein concentrations of approximately 400 mg/L culture. Based on the areas under the curve and the protein concentration of the fractionated material we are currently estimating yields of 50 mg protein for each isoform per liter of culture medium.

Example 3 - Delivery of Spike Proteins using Live Virus Vectors

A live virus approach using Modified Vaccinia virus Ankara (MVA) as a vector is now described. MVA has been proven to be extremely attenuated when compared to wild-type Vaccinia virus strain (Mayr et al., Infection 3:6-14, 1975; Werner et al., Arch.Virol. 64:247-256, 1980) and was established as exceptionally safe viral vector (Moss et al., *In* S. Cohen and A. Shafferman (eds.), Novel Strategies in the Design and Production of Vaccines, Plenum Press, New York, 1996, p. 7-13; Stittelaar et al., Vaccine 19:3700-3709, 2001; Sutter et al., Dev. Biol. Stand. 84:195-200, 1995). The following is a description of the construction of rMVAs expressing full-length recombinant SARS spike proteins. which can be used in vaccination methods against SARS, as described above.

For generating recombinant MVA a strategy called "transient dominant selection" (TDS) (Falkner et al., J. Virol. 64:3108-3111, 1990) can be used. The spike gene is amplified by PCR from a source clone (ACAM 250-0013; also see SEQ ID NO:38, Figure 65) and cloned into the BamHI-EcoRI sites of the insertion vector pTK53-gpt (Falkner et al., J. Virol. 64:3108-3111, 1990). The resulting plasmid, pTK-53-gpt-Spike (Figure 59), contains the Spike protein gene flanked by left (TK_L) and right (TK_R) shoulders of the vaccinia thymidine kinase (TK) gene, and is controlled by a powerful late Vaccinia P11 promoter. A schematic outline of the TDS approach is shown in Figure 60. When the resulting plasmid is transfected to the cells that have been infected with MVA virus, homologous recombination occurs due to homology between virus and plasmid TK gene sequences. As a result of a single crossover event, an unstable

intermediate virus containing the whole plasmid will be generated. Because of the presence of direct repeats, a second crossover event occurs and results in the formation of either wild type or recombinant virus containing the spike gene. All three types of genomes can be packaged separately into particles and are infectious, but only virus containing the *gpt* gene can form plaques under selective conditions.

Thus, the first two rounds of plaque isolation are done in presence of mycophenolic acid, xanthine, and hypoxanthine, which only allow the growth of viruses that express *E. coli gpt* (RM 2026 #7). The next two rounds are carried out without selection, and in the final plaque assay, the isolated virus can be checked for the expression of *gpt* and *tk* by the use selective media (RM 2026 #10). All gpt+tk- viruses should contain the spike gene, and this can be confirmed by PCR.

The viruses can be grown in chick embryo fibroblast (CEF) cells (Sutter et al., Proc. Natl. Acad. Sci. U.S.A 89:10847-10851, 1992) and/or baby hamster kidney (BHK) cells (Drexler et al., J. Gen. Virol. 79 (Pt 2):347-352, 1998). As an additional cell substrate for propagating MVA, a spontaneously immortalized chicken cell line, DF1, derived from 10 day old East Lancing Line (ELL-0) eggs (19) (ATCC # CRL-12203) can be used.

CEF, BHK, or DF-1 cells are infected with MVA as described in Gomez et al., Arch. Virol. 146:875-892, 2001. Briefly, 0.1 PFU/cell MVA or MVA recombinant in serum free medium can be used as infective dose. After 1 hour of virus adsorption, the inocula are removed and cells are supplemented with medium containing 2% serum and antibiotics. After 3-4 days of culture (depending on the type of cells), the cells are collected by centrifugation, washed and resuspended in medium, and sonicated; cell extracts are centrifuged at 2K/10 minutes, the supernatant collected, and the pellet resuspended in 1 mM of Na_2HPO_4, then re-extracted as described previously. Pooled supernatants are centrifuged 15K/30 minutes, the pellet resuspended by sonication in 1 mM of Na_2HPO_4, applied over 20-45 % (w/v) sucrose gradient in the same solution, and centrifuged at 15K/20 minutes. The virus band iscollected, diluted in 1 mM Na_2HPO_4, and sedimented at 15K/30 minutes; the virus pellet is then resuspended in a small volume of 1 mM Na_2HPO_4 and stored in aliquots at -70°C. To titrate or plaque purify the viruses

(Falkner et al., J. Virol. 62:1849-1854, 1988), a layer of semisolid medium (incubation medium + 1 % agarose) can be added to the infected cells. After incubation for one day at 37°C, a second layer of semisolid medium with 0.2 % of neutral red can be added, and after another 8-12 hours of incubation plaques are counted or collected by aspiration into glass pasteur pipettes. Virus can be released from agar by sonication or repetitive freezing-melting rounds.

An exemplary recombination protocol is described below. Note that although the term "cotransfection" is used in Figure 61, for the sake of simplicity, virus infection actually precedes plasmid transfection by 2-3 hours (Falkner et al., J. Virol. 64:3108-3111, 1990). After 2 hours of infection with MVA, CEF or DF-1 cells can be transfected with plasmid pTK53-gpt-spike and placed in gpt+ selective medium MXHAT (Boyle et al., Gene 65:123-128, 1988) (Dullbeco's modified Eagle Medium, 2.5 % fetal bovine serum, 25 μg of MPA per ml, 250 μg of xanthine per ml, 15 μg of hypoxanthine per ml). After 14 to 24 hours of incubation, plaques can be detected by staining with neutral red. Ten-twenty plaques of normal size and shape can be picked and then reassayed a second time under the same selective conditions. Three more plaque purification rounds are carried out under nonselective conditions. The *gpt*$^{+}$*tk*$^{-}$ phenotypes are then determined by plaque assay in the presence of MXHAT and 5-bromodeoxyuridine overlay as described by others (Chakrabarti et al., Mol. Cell Biol. 5:3403-3409, 1985; Mackett et al., J. Virol. 49:857-864, 1984). TK^{-} selection is then carried out as described previously.

As widespread smallpox vaccination is again considered (Abramson et al., Pediatrics 111:1431-1432, 2003), the prevalence of immunity to vaccinia virus could increase substantially. A possible preceding immunity to vaccinia virus could reduce its ability to serve as a vector for the delivery of recombinant genes used for other infectious deceases (Cooney et al., Lancet 337:567-572, 1991). An approach to alleviating this is by using either mucosal route (Belyakov et al., Proc. Natl. Acad. Sci. U.S.A 96:4512-4517, 1999) or DNA priming before vector boosting (Yang et al., J. Virol. 77:799-803, 2003).

The mucosal administration approach is based on the fact that migration of immune T cells between the mucosal and systemic immune systems is asymmetrically

restricted in the sense that cells traffic from mucosal system to the systemic system but not *vice versa*. Thus. systemic infection with vaccinia virus does not induce CTL that migrate to mucosal immune system, and apparently the virus does not infect mucosal tissues sufficiently under these circumstances to induce immunity (Belyakov et al., J. Clin. Invest 102:2072-2081, 1998; Belyakov et al., Proc. Natl. Acad. Sci. U.S.A 95:1709-1714, 1998; Belyakov et al., J. Virol. 72:8264-8272, 1998). On this basis, the mucosal system remains naïve to vaccinia virus. In contrast, mucosal infection with recombinant vaccinia virus induces not only CTL in the mucosa, but CTL that traffic out to the systemic immune system.

DNA priming has been shown to be highly effective in stimulating a primary immune response based on T-cell recognition of diverse subdominant epitopes (Barouch et al., J. Virol. 75:2462-2467, 2001). The response is presumably based on the ability of antigen-presenting cells to take up and present endogenously synthesized antigens. In the absence of proteins from viral vector, the primary immune response presumably can focus on the antigen of interest and facilitate the generation of memory T cells specific for the relevant antigen. Once these memory cells are present, viral vector proteins do not interfere with recall response, allowing a robust immune response to develop (Yang et al., J. Virol. 77:799-803, 2003).

Two rMVA constructs have been generated. rMVA-S and rMVA-N contain SARS spike (S) and nucleocapsid (N) genes, respectively. Expression cassettes carrying these target genes under the control of a late vaccinia virus promoter have been cloned into the thymidine kinase gene. Stable rMVA strains expressing both structural genes separately have been identified and the immunoreactivity of the expressed product determined by immunoblot with the polyclonal anti-SARS-CoV hyperimmune antibody (Figure 62). Both viruses have been plaque purified and amplified to a high titer. These viruses can be used, e.g., in the methods described above. As a specific example, they can be used in a prime boost strategy, in which they are administered mucosally in a priming step, which is followed by a parenteral boost with the recombinant protein. Other examples of regimens and routes that can be used are known in the art and discussed above.

All of the references cited herein are incorporated herein by reference in their entirety. Other embodiments are present in the following claims.

What is claimed is:

1. A vaccine for inducing an immune response to a human coronavirus that is the causative agent of Severe Acute Respiratory Syndrome (SARS) in a patient, said vaccine comprising a spike protein or a nucleocapsid protein of said virus, or an immunogenic fragment of either of these proteins, and a pharmaceutically acceptable carrier or diluent.

2. The vaccine of claim 1, wherein said immunogenic fragment of said spike protein comprises the S1 domain of said spike protein.

3. The vaccine of claim 2, wherein said immunogenic fragment of said spike protein further comprises the S2 domain of said spike protein, but not the coiled coil region of said spike protein.

4. The vaccine of claim 3, wherein said immunogenic fragment of said spike protein further comprises the coiled coil region of said spike protein.

5. The vaccine of claim 4, wherein said immunogenic fragment of said spike protein is in the form of a trimer.

6. The vaccine of claim 1, further comprising an adjuvant.

7. The vaccine of claim 6, wherein said adjuvant preferentially stimulates a Th1-type immune response.

8. The vaccine of claim 7, wherein said adjuvant is selected from the group consisting of an ISCOM, Ribi, DC-Chol, QS21, and MPL.

9. The vaccine of claim 6, wherein said adjuvant is alum.

10. The vaccine of claim 1, wherein said spike protein comprises an amino acid sequence that is substantially identical to the sequence of SEQ ID NO:37, or an immunogenic fragment thereof.

11. The vaccine of claim 1, wherein said nucleocapsid protein comprises an amino acid sequence that is substantially identical to the sequence of SEQ ID NO:35, or an immunogenic fragment thereof.

12. A vaccine for inducing an immune response to a human coronavirus that is the causative agent of Severe Acute Respiratory Syndrome (SARS) in a patient, said vaccine comprising a vector comprising a nucleic acid sequence encoding a spike protein or a nucleocapsid protein of said virus, or an immunogenic fragment of either of these proteins, and a pharmaceutically acceptable carrier or diluent.

13. The vaccine of claim 12, wherein said immunogenic fragment of said spike protein comprises the S1 domain of said spike protein.

14. The vaccine of claim 13, wherein said immunogenic fragment of said spike protein further comprises the S2 domain of said spike protein, but not the coiled coil region of said spike protein.

15. The vaccine of claim 14, wherein said immunogenic fragment of said spike protein further comprises the coiled coil region of said spike protein.

16. The vaccine of claim 12, wherein said vector is a viral vector.

17. The vaccine of claim 16, wherein said vector comprises a poxvirus or an adenovirus.

18. The vaccine of claim 17, wherein said poxvirus is a modified vaccinia akara virus.

19. A method for producing a spike protein or a nucleocapsid protein of a human coronavirus, or an immunogenic fragment thereof, said method comprising introducing into cells a vector comprising a nucleic acid sequence encoding said protein or said fragment, under conditions in which said protein or fragment is expressed in said cells.

20. The method of claim 19, wherein said immunogenic fragment of said spike protein comprises the S1' domain of said spike protein.

21. The method of claim 20, wherein said immunogenic fragment of said spike protein further comprises the S2 domain of said spike protein, but not the coiled coil region of said spike protein.

22. The method of claim 21, wherein said immunogenic fragment of said spike protein further comprises the coiled coil region of said spike protein.

23. The method of claim 19, wherein said cells are yeast cells, mammalian cells, insect cells, or bacterial cells.

24. The method of claim 23, wherein said yeast cells are *Pichia pastoris* cells.

25. A method of inducing an immune response to a human coronavirus that is the causative agent of Severe Acute Respiratory Syndrome (SARS) in a patient, said method comprising administering the vaccine of claim 1 or claim 12 to said patient.

26. A substantially pure spike protein of a human coronavirus that is the causative agent of Severe Acute Respiratory Syndrome (SARS). or an immunogenic fragment thereof.

27. The protein of claim 26. wherein said immunogenic fragment of said spike protein comprises S1 domain of said spike protein.

28. The protein of claim 27, wherein said immunogenic fragment of said spike protein further comprises the S2 domain of said spike protein, but not the coiled coil region of said spike protein.

29. The protein of claim 28, wherein said immunogenic fragment of said spike protein further comprises the coiled coil region of said spike protein.

30. The protein of claim 26, wherein said spike protein or fragment comprises a sequence that is substantially identical to the sequence of SEQ ID NO:37, or a fragment thereof.

31. The protein of claim 26. wherein said spike protein or fragment comprises the sequence of SEQ ID NO:37, or a fragment thereof.

32. The protein of claim 26, wherein said spike protein or fragment is in the form of a trimer.

33. An isolated nucleic acid molecule encoding a spike protein of a human coronavirus or an immunogenic fragment thereof.

34. The nucleic acid molecule of claim 33, wherein said immunogenic fragment of said spike protein comprises the S1 domain of said spike protein.

35. The nucleic acid molecule of claim 34, wherein said immunogenic fragment of said spike protein further comprises the S2 domain of said spike protein, but not the coiled coil region of said spike protein.

36. The nucleic acid molecule of claim 35, wherein said immunogenic fragment of said spike protein further comprises the coiled coil region of said spike protein.

37. The nucleic acid molecule of claim 33, wherein said nucleic acid molecule comprises the sequence of SEQ ID NO:36.

38. The nucleic acid molecule of claim 33, wherein said nucleic acid molecule hybridizes to the complement of the sequence of SEQ ID NO:36 under highly stringent conditions.

39. A nucleic acid molecule probe comprising a sequence that hybridizes to the sequence of SEQ ID NO:36 or the complement thereof under highly stringent conditions.

40. An antibody that specifically binds to the protein or immunogenic fragment of claim 26.

41. A substantially pure nucleocapsid protein of a human coronavirus that is the causative agent of Severe Acute Respiratory Syndrome (SARS), or an immunogenic fragment thereof.

42. The protein of claim 41, wherein said spike protein or fragment comprises a sequence that is substantially identical to the sequence of SEQ ID NO:37, or a fragment thereof.

43. The protein of claim 41, wherein said spike protein or fragment comprises the sequence of SEQ ID NO:37, or a fragment thereof.

44. An isolated nucleic acid molecule encoding a nucleocapsid protein of a human coronavirus or an immunogenic fragment thereof.

45. The nucleic acid molecule of claim 44, wherein said nucleic acid molecule comprises the sequence of SEQ ID NO:34.

46. The nucleic acid molecule of claim 44, wherein said nucleic acid molecule hybridizes to the complement of the sequence of SEQ ID NO:34 under highly stringent conditions.

47. A nucleic acid molecule probe comprising a sequence that hybridizes to the sequence of SEQ ID NO:34 or the complement thereof under highly stringent conditions.

48. An antibody that specifically binds to the protein or immunogenic fragment of claim 41.

49. A method of preventing or treating infection by a human coronavirus that is the causative agent of Severe Acute Respiratory Syndrome (SARS) in a patient, comprising administering to the patient the antibody of claim 40 or claim 48.

50. The method of claim 49, wherein said antibody is a polyclonal hyperimmune globulin preparation.

Figure 1

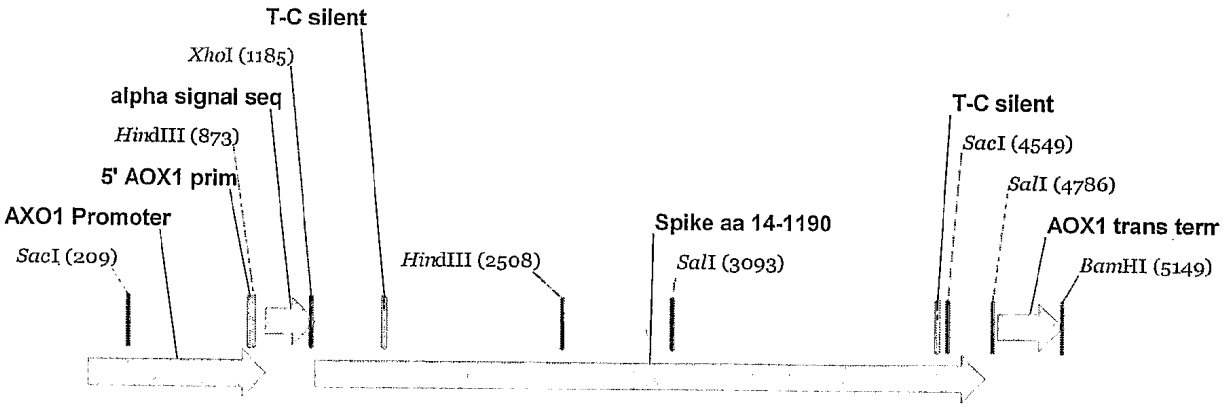
pPICZ alpha 1190 clone P5-12
Deduced Amino Acid Sequence (SEQ ID NO:1)

Alpha Signal Amino Acids -
mrfpsiftavlfaassalaapvnttttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike Amino Acids (14 thru 1190, does not contain first 13 amino acid leader).
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfytttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvdcnmyicgdstecanlllqygs
fctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnfsqilpdplkptkrsfiedllfnkvtladagfmkqyge
clgdinardlicaqkfngltvlppllтddmiaaytaalvsgtatagwtfgagaalqipfamqmayrfngigvtqnvlyen
qkqianqfnkaisqiqesltttstalgklqdvvnqnaqalntlvkqlssnfgaissvlndilsrldkveaevqidrlitgrlqsl
qtyvtqqliraaeirasanlaatkmsecvlgqskrvdfcgkgyhlmsfpqaaphgvvflhvtyvpsqernfttapaiche
gkayfpregvfvfngtswfitqrnffspqiittdntfvsgncdvvigiinntvydplqpeldsfkeeldkyfknhtspdvd
lgdisginasvvniqkeidrlnevaknlneslidlqelgkyeq*

Figure 2

pPICZ alpha 1190 clone P5-12
Map of AXO1 Promoter - alpha – Spike – Term.



P5-12 seq
5153 bp

Figure 3

pPICZ alpha 1190 clone P5-12 map

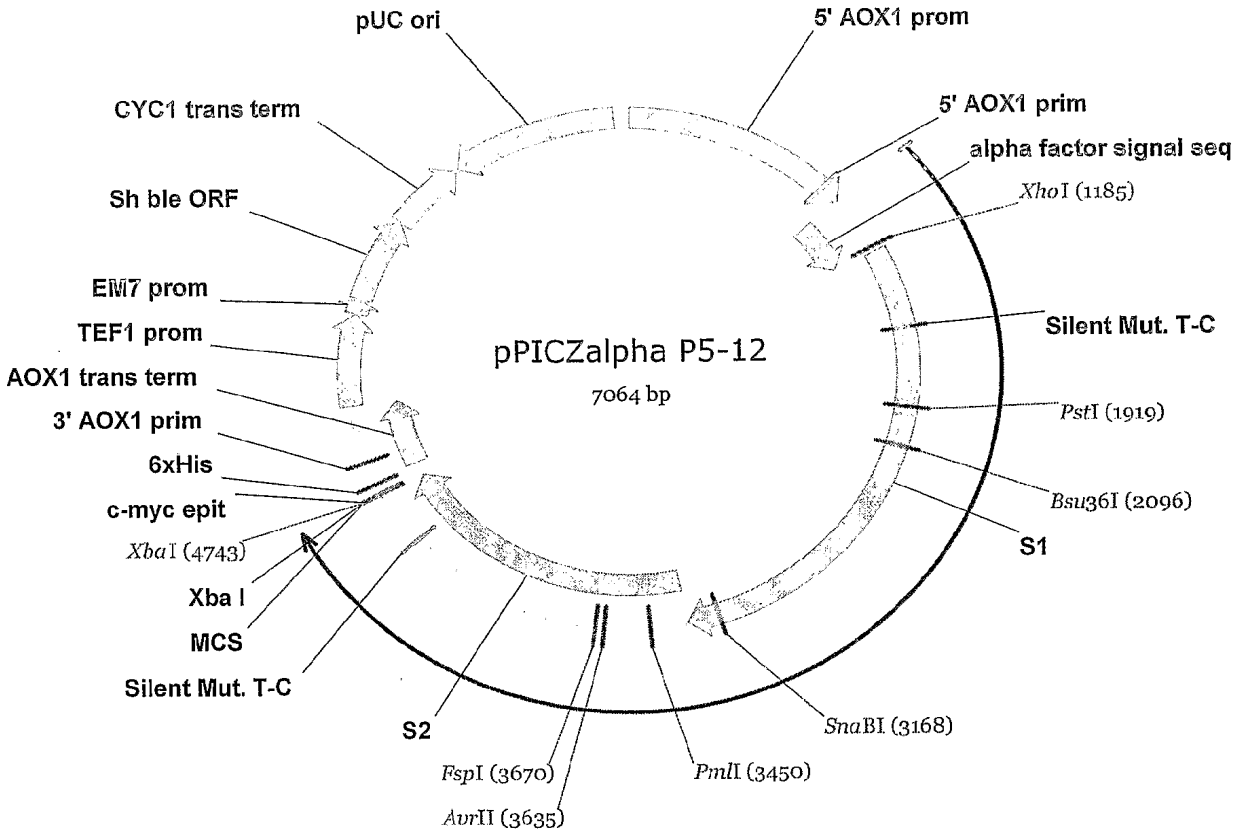


Figure 4

pPICZ alpha 1190 clone P5-12 sequence, from linear map (SEQ ID NO:2)

- AXO1 Promoter – base 1 to 940

agatctaacatccaaagacgaaaggttgaatgaaacctttttgccatccgacatccacaggtccattctcacacataagtgcc
aaacgcaacaggaggggatacactagcagcagaccgttgcaaacgcaggacctccactcctcttctcctcaacacccact
tttgccatcgaaaaaccagcccagttattgggcttgattggagctcgctcattccaattccttctattaggctactaacaccatg
actttattagcctgtctatcctggcccccctggcgaggttcatgtttgtttatttccgaatgcaacaagctccgcattacacccga
acatcactccagatgagggctttctgagtgtggggtcaaatagtttcatgttccccaaatggcccaaaactgacagtttaaac
gctgtcttggaacctaatatgacaaaagcgtgatctcatccaagatgaactaagtttggttcgttgaaatgctaacggccagtt
ggtcaaaaagaaacttccaaaagtcggcataccgtttgtcttgtttggtattgattgacgaatgctcaaaaataatctcattaatg
cttagcgcagtctctctatcgcttctgaaccccggtgcacctgtgccgaaacgcaaatggggaaacacccgcttttttggatg
attatgcattgtctccacattgtatgcttccaagattctggtgggaatactgctgatagcctaacgttcatgatcaaaatttaactg
ttctaacccctacttgacagcaatatataaacagaaggaagctgccctgtcttaaacctttttttttatcatcattattagcttacttt
cataattgcgactggttccaattgacaagcttttgattttaacgacttttaacgacaacttgagaagatcaaaaaacaactaatta
ttcgaaacg

- alpha Signal Sequence – base 941 to 1207

atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaag
agaggctgaagct

- Spike 1190 – base 1208 to 4741

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttaaggatggtatttatttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgcgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgccttttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtatatctgcggagattctactgaatgtgctaatttgcttctccaatatgg
cagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaag

tcaaacaaatgtacaaaaccccaactttgaaatattttggtggttttaatttttcacaaatattacctgaccctctaaagccaacta
agaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgcctag
gtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatgatt
gctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaataccttttg
ctatgcaaatggcatataggttcaatggcattggagttacccaaaatgttctctatgagaaccaaaaacaaatcgccaaccaa
tttaacaaggcgattagtcaaattcaagaatcacttacaacaacatcaactgcattgggcaagctgcaagacgttgttaacca
gaatgctcaagcattaaacacacttgttaaacaacttagctctaattttggtgcaatttcaagtgtgctaaatgatatcctttcgcg
acttgataaagtcgaggcggaggtacaaattgacaggttaattacaggcagacttcaaagccttcaaacctatgtaacacaa
caactaatcagggctgctgaaatcagggcttctgctaatcttgctgctactaaaatgtctgagtgtgttcttggacaatcaaaaa
gagttgacttttgtggaaagggctaccaccttatgtccttcccacaagcagccccgcatggtgttgtcttcctacatgtcacgt
atgtgccatcccaggagaggaacttcaccacagcgccagcaatttgtcatgaaggcaaagcatacttccctcgtgaaggtg
tttttgtgtttaatggcacttcttggtttattacacagaggaacttcttttctccacaaataattactacagacaatacatttgtctca
ggaaattgcgatgtcgttattggcatcattaacaacacagtttatgatcctctgcaacctgagctcgactcattcaaagaagag
ctggacaagtacttcaaaaatcatacatcaccagatgttgatcttggcgacatttcaggcattaacgcttctgtcgtcaacattc
aaaaagaaattgaccgcctcaatgaggtcgctaaaaatttaaatgaatcactcattgaccttcaagaattgggaaaatatgag
caataa

- MCS/Xba I/c-myc epit/6xHis/3'AOX1 prim – base 4742 to 4811
tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattga

- AOX1 Transcription Terminator - base 4812 to 5153
gtttgtagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagagg
atgtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcat
tttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattc
gagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

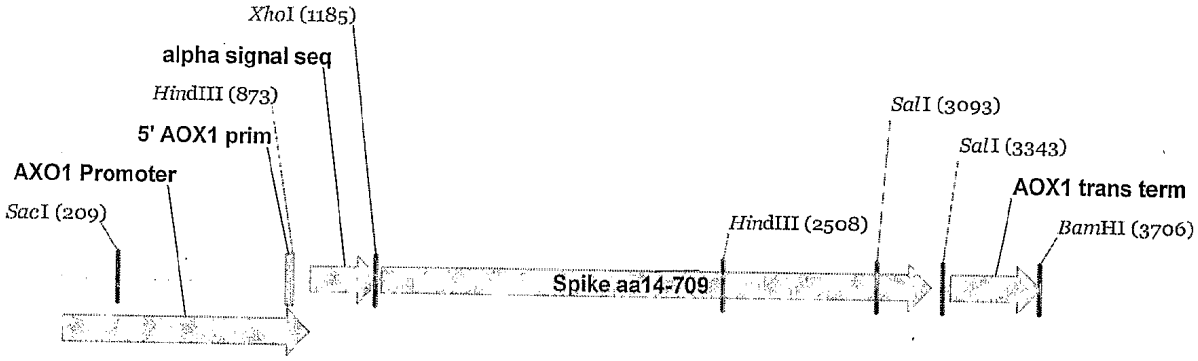
Figure 5

pPICZ alpha 709 clone P1-2
Deduced Amino Acid Sequence (SEQ ID NO:3)

Alpha Signal Amino Acids -
mrfpsiftavlfaassalaapvntttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike Amino Acids (14 thru 709, does not contain first 13 amino acid leader).
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfytttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevm*
pPICZ alpha 709 clone P1-2
Linear Map

Figure 6



P1-2 709
3710 bp

Figure 7

pPICZ alpha 709 clone P1-2
Sequence, from linear map (SEQ ID NO:4)

- AXO1 Promoter – base 1 to 940

agatctaacatccaaagacgaaaggttgaatgaaacctttttgccatccgacatccacaggtccattctcacacataagtgcc
aaacgcaacaggaggggatacactagcagcagaccgttgcaaacgcaggacctccactcctcttctcctcaacacccact
tttgccatcgaaaaaccagcccagttattgggcttgattggagctcgctcattccaattccttctattaggctactaacaccatg
actttattagcctgtctatcctggcccccctggcgaggttcatgtttgtttatttccgaatgcaacaagctccgcattacacccga
acatcactccagatgagggctttctgagtgtggggtcaaatagtttcatgttccccaaatggcccaaaactgacagtttaaac
gctgtcttggaacctaatatgacaaaagcgtgatctcatccaagatgaactaagtttggttcgttgaaatgctaacggccagtt
ggtcaaaaagaaacttccaaaagtcggcataccgtttgtcttgtttggtattgattgacgaatgctcaaaaataatctcattaatg
cttagcgcagtctctctatcgcttctgaaccccggtgcacctgtgccgaaacgcaaatggggaaacacccgctttttggatg
attatgcattgtctccacattgtatgcttccaagattctggtgggaatactgctgatagcctaacgttcatgatcaaaatttaactg
ttctaaccccctacttgacagcaatatataaacagaaggaagctgccctgtcttaaacctttttttttatcatcattattagcttacttt
cataattgcgactggttccaattgacaagcttttgattttaacgacttttaacgacaacttgagaagatcaaaaaacaactaatta
ttcgaaacg

- alpha Signal Sequence – base 941 to 1207

atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaggggtatctctcgagaaaag
agaggctgaagct

- Spike – base 1208 to 3298

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttttaaggatggtatttatttttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgcgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaatttttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactataa

- MCS/Xba I/c-myc epit/6xHis/3'AOX1 prim – base 3299 to 3368
tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattga

- AOX1 Transcription Terminator – base 3369 to 3710
gtttgtagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagagg
atgtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcat
tttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattc
gagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

Figure 8

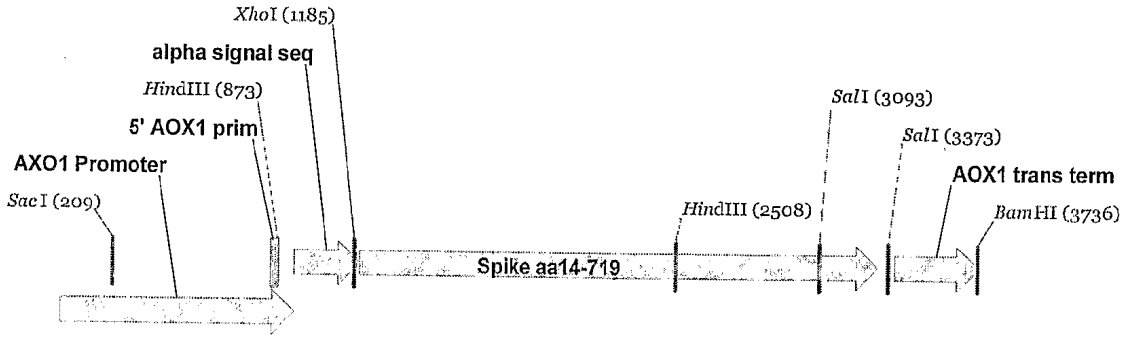
pPICZ alpha 719 clone P1-2
Deduced Amino Acid Sequence (SEQ ID NO:5)

Alpha Signal Amino Acids -
mrfpsiftavlfaassalaapvnttttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike Amino Acids (14 thru 719, does not contain first 13 amino acid leader).
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfytttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvd*

Figure 9

pPICZ alpha 719 clone P1-2
Linear Map



P1-2 719 linear map
3740 bp

Figure 10

pPICZ alpha 719 clone P1-2
Sequence, from linear map (SEQ ID NO:6)

- AXO1 Promoter – base 1 to 940

agatctaacatccaaagacgaaaggttgaatgaaacctttttgccatccgacatccacaggtccattctcacacataagtgcc
aaacgcaacaggaggggatacactagcagcagaccgttgcaaacgcaggacctccactcctcttctcctcaacacccact
tttgccatcgaaaaaccagcccagttattgggcttgattggagctcgctcattccaattccttctattaggctactaacaccatg
actttattagcctgtctatcctggcccccctggcgaggttcatgtttgtttatttccgaatgcaacaagctccgcattacacccga
acatcactccagatgagggctttctgagtgtggggtcaaatagtttcatgttccccaaatggcccaaaactgacagtttaaac
gctgtcttggaacctaatatgacaaaagcgtgatctcatccaagatgaactaagtttggttcgttgaaatgctaacggccagtt
ggtcaaaaagaaacttccaaaagtcggcataccgtttgtcttgtttggtattgattgacgaatgctcaaaaataatctcattaatg
cttagcgcagtctctctatcgcttctgaaccccggtgcacctgtgccgaaacgcaaatggggaaacacccgctttttggatg
attatgcattgtctccacattgtatgcttccaagattctggtgggaatactgctgatagcctaacgttcatgatcaaaatttaactg
ttctaacccctacttgacagcaatatataaacagaaggaagctgccctgtcttaaacctttttttttatcatcattattagcttacttt
cataattgcgactggttccaattgacaagcttttgattttaacgacttttaacgacaacttgagaagatcaaaaaacaactaatta
ttcgaaacg

- alpha Signal Sequence – base 941 to 1207

atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaggggtatctctcgagaaaag
agaggctgaagct

- Spike – base 1208 to 3328

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcataccttttaaggatggtatttattttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgcgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattaa

- MCS/Xba I/c-myc epit/6xHis/3'AOX1 prim – base 3329 to 3398
tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattga

- AOX1 Transcription Terminator – base 3399 to 3740
gtttgtagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagagg
atgtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcat
tttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattc
gagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

Figure 11

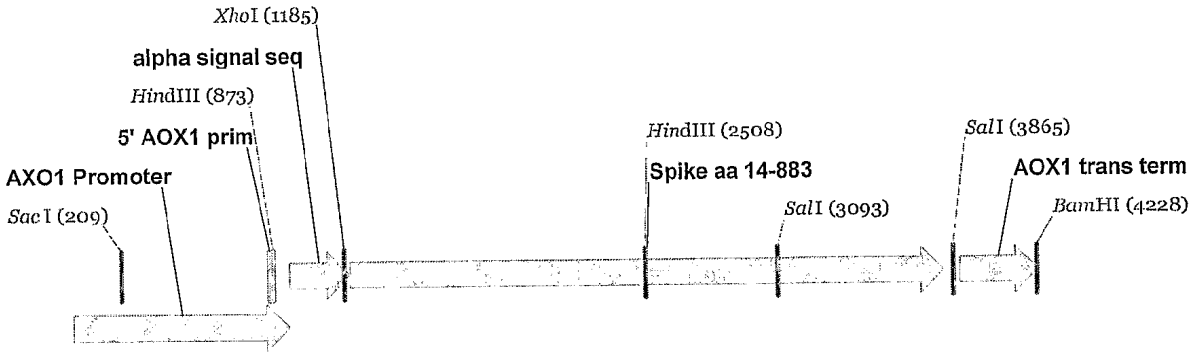
pPICZ alpha 883 clone P3-10
Deduced Amino Acid Sequence (SEQ ID NO:7)

Alpha Signal Amino Acids -
mrfpsiftavlfaassalaapvnttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike Amino Acids (14 thru 883, does not contain first 13 amino acid leader).
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfyttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvdcnmyicgdstecanlllqygs
fctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnfsqilpdplkptkrsfiedllfnkvtladagfmkqyge
clgdinardlicaqkfngltvlppllddmiaaytaalvsgtatagwtfgagaalqipfamq*

Figure 12

pPICZ alpha 883 clone P3-10
Linear Map



P3-10 883
4232 bp

Figure 13

pPICZ alpha 883 clone P3-10
Sequence, from linear map (SEQ ID NO:8)

- AXO1 Promoter – base 1 to 940

agatctaacatccaaagacgaaaggttgaatgaaacctttttgccatccgacatccacaggtccattctcacacataagtgcc
aaacgcaacaggaggggatacactagcagcagaccgttgcaaacgcaggacctccactcctcttctcctcaacacccact
tttgccatcgaaaaaccagcccagttattgggcttgattggagctcgctcattccaattccttctattaggctactaacaccatg
actttattagcctgtctatcctggcccccctggcgaggttcatgtttgtttatttccgaatgcaacaagctccgcattacacccga
acatcactccagatgagggctttctgagtgtggggtcaaatagtttcatgttccccaaatggcccaaaactgacagtttaaac
gctgtcttggaacctaatatgacaaaagcgtgatctcatccaagatgaactaagtttggttcgttgaaatgctaacggccagtt
ggtcaaaaagaaacttccaaaagtcggcataccgtttgtcttgtttggtattgattgacgaatgctcaaaaataatctcattaatg
cttagcgcagtctctctatcgcttctgaaccccggtgcacctgtgccgaaacgcaaatggggaaacacccgcttttttggatg
attatgcattgtctccacattgtatgcttccaagattctggtgggaatactgctgatagcctaacgttcatgatcaaaatttaactg
ttctaaccccctacttgacagcaatatataaacagaaggaagctgccctgtcttaaaccttttttttttatcatcattattagcttacttt
cataattgcgactggttccaattgacaagcttttgattttaacgacttttaacgacaacttgagaagatcaaaaaacaactaatta
ttcgaaacg

- alpha Signal Sequence – base 941 to 1207

atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaag
agaggctgaagct

- Spike – base 1208 to 3820

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttaaggatggtatttatttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgcgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgccttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtacatctgcggagattctactgaatgtgctaatttgcttctccaatatg

gtagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaa
gtcaaacaaatgtacaaaaccccaactttgaaatattttggtggttttaatttttcacaaatattacctgaccctctaaagccaact
aagaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgccta
ggtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatga
ttgctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaatacctttt
gctatgcaataa

- MCS/Xba I/c-myc epit/6xHis/3'AOX1 prim – base 3821 to 3890
tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattga

- AOX1 Transcription Terminator – base 3891 to 4232
gtttgtagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagagg
atgtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcat
tttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattc
gagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

Figure 14

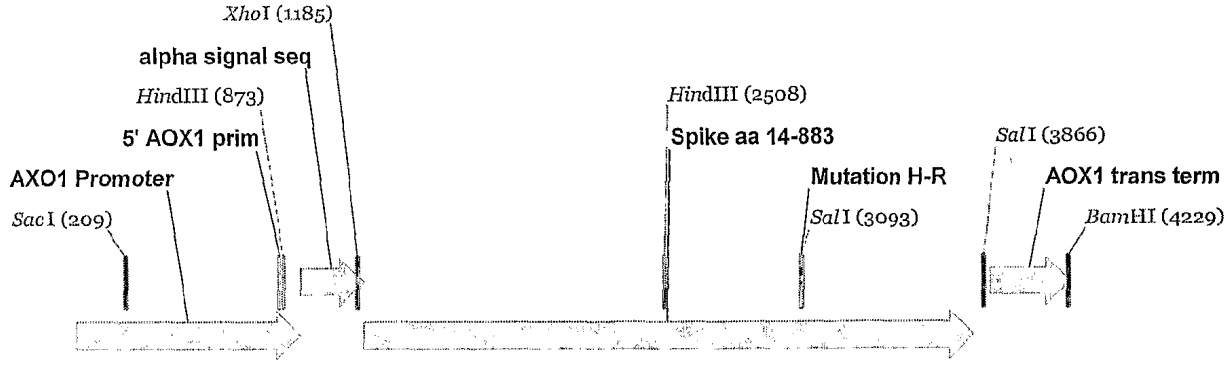
pPICZ alpha 883m clone P3-10
Deduced Amino Acid Sequence (SEQ ID NO:9)

Alpha Signal Amino Acids -
mrfpsiftavlfaassalaapvnttttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike Amino Acids (14 thru 883, does not contain first 13 amino acid leader).
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfyttigigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaervdtsyecdipigagicas
yhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvdcnmyicgdstecanlllqygsf
ctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnfsqilpdplkptkrsfiedllfnkvtladagfmkqygec
lgdinardlicaqkfngltvlppllttddmiaaytaalvsgtatagwtfgagaalqipfamq*

Figure 15

pPICZ alpha 883m clone P3-10
Linear Map



P3-10 883m

4233 bp

Figure 16

pPICZ alpha 883m clone P3-10
Sequence, from linear map (SEQ ID NO:10)

- AXO1 Promoter – base 1 to 940

agatctaacatccaaagacgaaaggttgaatgaaacctttttgccatccgacatccacaggtccattctcacacataagtgcc
aaacgcaacaggaggggatacactagcagcagaccgttgcaaacgcaggacctccactcctcttctcctcaacacccact
tttgccatcgaaaaaccagcccagttattgggcttgattggagctcgctcattccaattccttctattaggctactaacaccatg
actttattagcctgtctatcctggcccccctggcgaggttcatgtttgtttatttccgaatgcaacaagctccgcattacacccga
acatcactccagatgagggctttctgagtgtggggtcaaatagtttcatgttccccaaatggcccaaaactgacagtttaaac
gctgtcttggaacctaatatgacaaaagcgtgatctcatccaagatgaactaagtttggttcgttgaaatgctaacggccagtt
ggtcaaaaagaaacttccaaaagtcggcataccgtttgtcttgtttggtattgattgacgaatgctcaaaaataatctcattaatg
cttagcgcagtctctctatcgcttctgaaccccggtgcacctgtgccgaaacgcaaatggggaaacacccgctttttggatg
attatgcattgtctccacattgtatgcttccaagattctggtgggaatactgctgatagcctaacgttcatgatcaaaatttaactg
ttctaacccctacttgacagcaatatataaacagaaggaagctgccctgtcttaaacctttttttttatcatcattattagcttacttt
cataattgcgactggttccaattgacaagcttttgatttaacgacttttaacgacaacttgagaagatcaaaaaacaactaatta
ttcgaaacg

- alpha Signal Sequence – base 941 to 1207

atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaag
agaggctgaagct

- Spike – base 1208 to 3820

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccatttttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttttaaggatggtatttatttttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgcgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcgtgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtacatctgcggagattctactgaatgtgctaatttgcttctccaatatg

gtagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaa
gtcaaacaaatgtacaaaaccccaactttgaaatattttggtggttttaatttttcacaaatattacctgaccctctaaagccaact
aagaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgccta
ggtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatga
ttgctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaatacctttt
gctatgcaataa

- MCS/Xba I/c-myc epit/6xHis/3'AOX1 prim – base 3821 to 3890

tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattga

- AOX1 Transcription Terminator – base 3891 to 4232

gtttgtagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagagg
atgtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcat
tttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattc
gagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

Figure 17

pGAPZ alpha 1190 clone G5-14 map

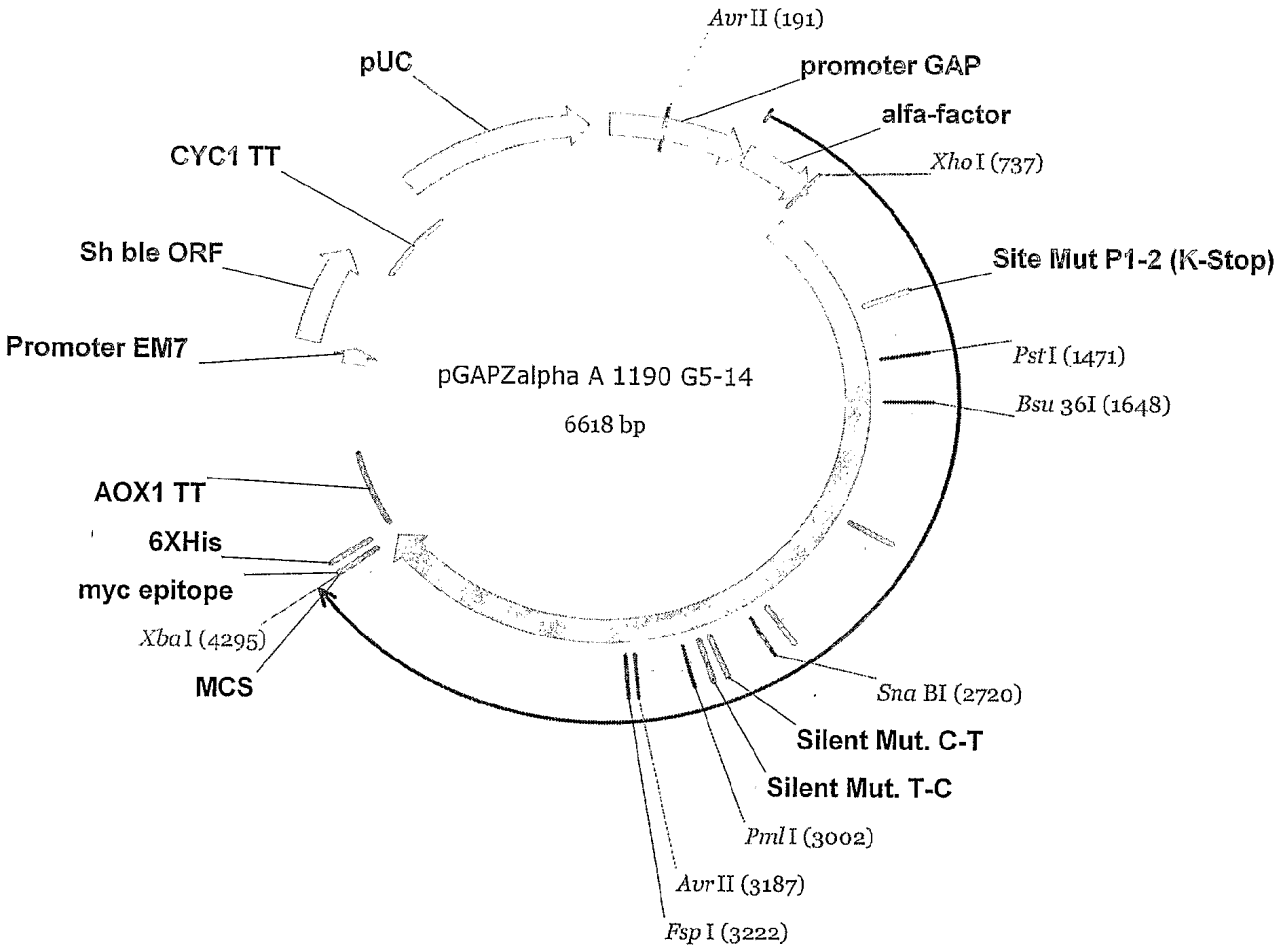


Figure 18

PGAPZ alpha 1190 clone G5-14
Deduced Amino Acid Sequence (SEQ ID NO:11)

Alpha Signal Sequence -
mrfpsiftavlfaassalaapvnttttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike amino acids 14 to 1190, does not include the first 13 amino acid leader
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfytttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvdcnmyicgdstecanlllqygs
fctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnfsqilpdplkptkrsfiedllfnkvtladagfmkqyge
clgdinardlicaqkfngltvlppllddmiaaytaalvsgtatagwtfgagaalqipfamqmayrfngigvtqnvlyen
qkqianqfnkaisqiqesltttstalgklqdvvnqnaqalntlvkqlssnfgaissvlndilsrldkveaevqidrlitgrlqsl
qtyvtqqliraaeirasanlaatkmsecvlgqskrvdfcgkgyhlmsfpqaaphgvvflhvtyvpsqernfttapaiche
gkayfpregvfvfngtswfitqrnffspqiittdntfvsgncdvvigiinntvydplqpeldsfkeeldkyfknhtspdvd
lgdisginasvvniqkeidrlnevaknlneslidlqelgkyeq*

Figure 19

PGAPZ alpha 1190 clone G5-14
Linear Map

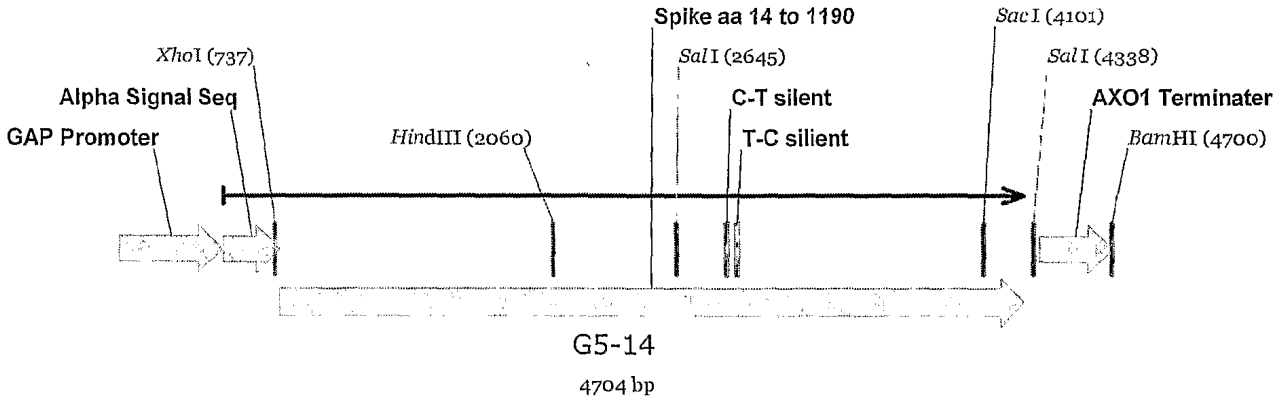


Figure 20

PGAPZ alpha 1190 Clone G5-14
Sequence (SEQ ID NO:12)

GAP Pomoter – base 1 to 483

agatcttttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctggctccgttgcaactccgaacg
acctgctggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaacgtctcttcccttctctctcct
tccaccgcccgttaccgtccctaggaaattttactctgctggagagcttcttctacggccccccttgcagcaatgctcttcccag
cattacgttgcgggtaaaacggaggtcgtgtacccgacctagcagcccagggatggaaaagtcccggccgtcgctggca
ataatagcgggcggacgcatgtcatgagattattggaaaccaccagaatcgaatataaaaggcgaacacctttcccaatttt
ggtttctcctgacccaaagactttaaatttaatttatttgtccctatttcaatcaattgaacaactat

-Spacer - base 484 to 492

ttcgaaacg

- Alpha Signal Sequence – base 493 to 759

atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaggggtatctctcgagaaaag
agaggctgaagct

- Spike aa 14 to 1190 – base 760 to 4293

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcataccttttaaggatggtatttattttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagcctttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtatatctgcggagattctactgaatgtgctaatttgcttctccaatatgg
cagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaag
tcaaacaaatgtacaaaaccccaactttgaaatattttggcggttttaatttttcacaaatattacctgaccctctaaagccaacta
agaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgcctag

gtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatgatt
gctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaatacctttg
ctatgcaaatggcatataggttcaatggcattggagttacccaaaatgttctctatgagaaccaaaaacaaatcgccaaccaa
tttaacaaggcgattagtcaaattcaagaatcacttacaacaacatcaactgcattgggcaagctgcaagacgttgttaacca
gaatgctcaagcattaaacacacttgttaaacaacttagctctaattttggtgcaatttcaagtgtgctaaatgatatcctttcgcg
acttgataaagtcgaggcggaggtacaaattgacaggttaattacaggcagacttcaaagccttcaaacctatgtaacacaa
caactaatcagggctgctgaaatcagggcttctgctaatcttgctgctactaaaatgtctgagtgtgttcttggacaatcaaaaa
gagttgacttttgtggaaagggctaccaccttatgtccttcccacaagcagccccgcatggtgttgtcttcctacatgtcacgt
atgtgccatcccaggagaggaacttcaccacagcgccagcaatttgtcatgaaggcaaagcatacttccctcgtgaaggtg
tttttgtgtttaatggcacttcttggtttattacacagaggaacttcttttctccacaaataattactacagacaatacatttgtctca
ggaaattgtgatgtcgttattggcatcattaacaacacagtttatgatcctctgcaacctgagctcgactcattcaaagaagag
ctggacaagtacttcaaaaatcatacatcaccagatgttgatcttggcgacatttcaggcattaacgcttctgtcgtcaacattc
aaaaagaaattgaccgcctcaatgaggtcgctaaaaatttaaatgaatcactcattgaccttcaagaattgggaaaatatgag
caataa

- MCS... base 4294 to 4363

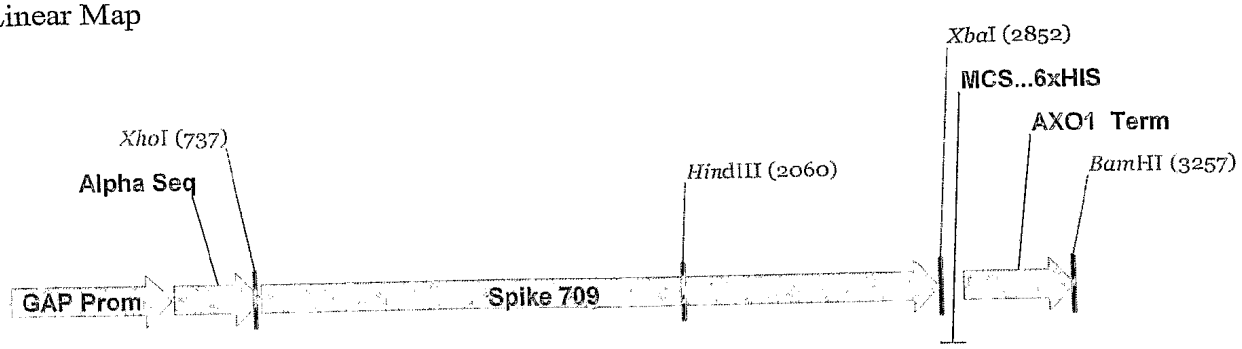
tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattga

- AXO1 Terminater base 4364 to 4704

gttttagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggat
gtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcatttt
gtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattcga
gtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

Figure 21

pGAPZ alpha 709 clone G1-8
Linear Map



709 G1-8 linear map

3261 bp

Figure 22

PGAPZ alpha 709 Clone G1-8
Sequence (SEQ ID NO:13)

GAP Pomoter – base 1 to 483
agatctttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctggctccgttgcaactccgaacg
acctgctggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaacgtctcttcccttctctctcct
tccaccgcccgttaccgtccctaggaaattttactctgctggagagcttcttctacggccccccttgcagcaatgctcttcccag
cattacgttgcgggtaaaacggaggtcgtgtacccgacctagcagcccagggatggaaaagtcccggccgtcgctggca
ataatagcgggcggacgcatgtcatgagattattggaaaccaccagaatcgaatataaaaggcgaacacctttcccaatttt
ggtttctcctgacccaaagactttaaatttaatttatttgtccctatttcaatcaattgaacaactat

-Spacer - base 484 to 492
ttcgaaacg

- Alpha Signal Sequence – base 493 to 759
atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaag
agaggctgaagct

- Spike – base 760 to 2847
agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccatttttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcataccttttaaggatggtatttatttttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaatttttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagcctttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt

ccettctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaactttcaattagcattactacagaagtaatgta
a

- MCS...6xHIS base 2848 to 2920
tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattga

- AXO1 Terminater base 2921 to 3261
gttttagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggat
gtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcatttt
gtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattcga
gtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

Figure 23

PGAPZ alpha 709 clone G1-8
Deduced Amino Acid Sequence (SEQ ID NO:14)

Alpha Signal Sequence -
mrfpsiftavlfaassalaapvntttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike amino acids 14 to 709, does not include the first 13 amino acid leader
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfytttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevm*

Figure 24

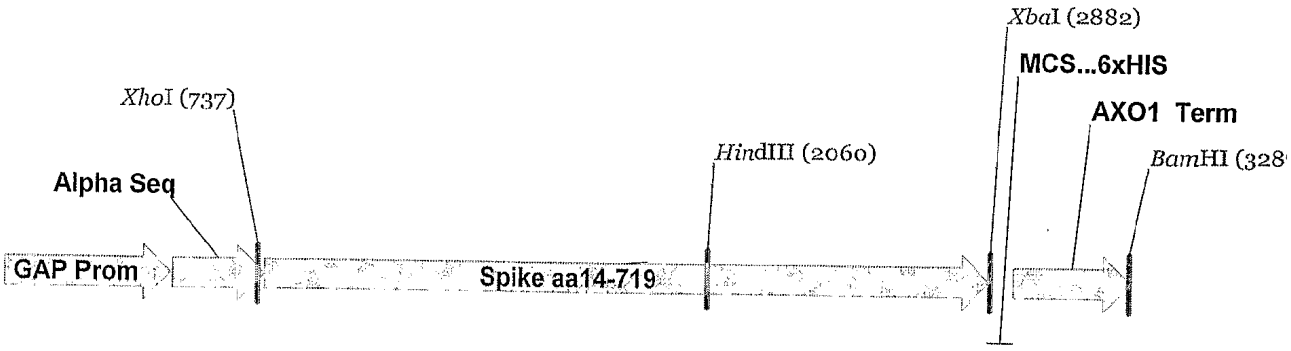
PGAPZ alpha 719 clone G1-8
Deduced Amino Acid Sequence (SEQ ID NO:15)

Alpha Signal Sequence -
mrfpsiftavlfaassalaapvnttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike amino acids 14 to 719, does not include the first 13 amino acid leader
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfytttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvd*

Figure 25

pGAPZ alpha 719 clone G1-8
Linear Map



719 G1-8 linear map

3291 bp

Figure 26

PGAPZ alpha 719 Clone G1-8
Sequence (SEQ ID NO:16)

GAP Pomoter – base 1 to 483

agatctttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctggctccgttgcaactccgaacg
acctgctggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaacgtctcttcccttctctctcct
tccaccgcccgttaccgtccctaggaaattttactctgctggagagcttcttctacggccccccttgcagcaatgctcttcccag
cattacgttgcgggtaaaacggaggtcgtgtacccgacctagcagcccagggatggaaaagtcccggccgtcgctggca
ataatagcgggcggacgcatgtcatgagattattggaaaccaccagaatcgaatataaaaggcgaacacctttcccaatttt
ggtttctcctgacccaaagactttaaatttaatttatttgtccctatttcaatcaattgaacaactat

-Spacer - base 484 to 492

ttcgaaacg

- Alpha Signal Sequence – base 493 to 759

atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccatttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaag
agaggctgaagct

- Spike – base 760 to 2877

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttttaaggatggtatttatttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgccttttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattaa

- MCS...6xHIS base 2878 to 2950

tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcatcattga

- AXO1 Terminater base 2951 to 3291

gttttagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggat
gtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcatttt
gtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattcga
gtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

Figure 27

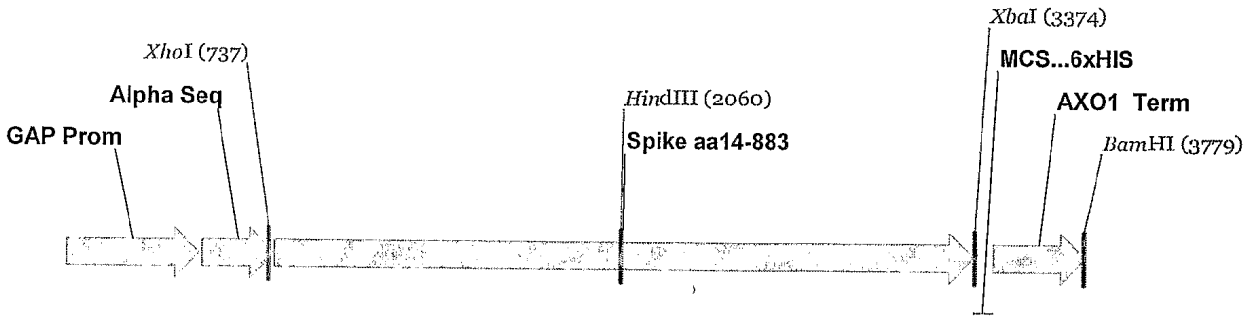
PGAPZ alpha 883 clone G3-7
Deduced Amino Acid Sequence (SEQ ID NO:17)

Alpha Signal Sequence -
mrfpsiftavlfaassalaapvnttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike amino acids 14 to 883, does not include the first 13 amino acid leader
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfytttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvdcnmyicgdstecanlllqygs
fctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnfsqilpdplkptkrsfiedllfnkvtladagfmkqyge
clgdinardlicaqkfngltvlppllddmiaaytaalvsgtatagwtfgagaalqipfamq*

Figure 28

pGAPZ alpha 883 clone G3-7
Linear Map



883 G3-7 linear map
3783 bp

Figure 29

PGAPZ alpha 883 Clone G3-7
Sequence (SEQ ID NO:18)

GAP Pomoter – base 1 to 483

agatctttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctggctccgttgcaactccgaacg
acctgctggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaacgtctcttcccttctctctcct
tccaccgcccgttaccgtccctaggaaattttactctgctggagagcttcttctacggccccccttgcagcaatgctcttcccag
cattacgttgcgggtaaaacggaggtcgtgtacccgacctagcagcccagggatggaaaagtcccggccgtcgctggca
ataatagcgggcggacgcatgtcatgagattattggaaaccaccagaatcgaatataaaaggcgaacacctttcccaatttt
ggtttctcctgacccaaagactttaaatttaatttatttgtccctatttcaatcaattgaacaactat

-Spacer - base 484 to 492

ttcgaaacg

- Alpha Signal Sequence – base 493 to 759

atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaggggtatctctcgagaaaag
agaggctgaagct

- Spike aa 14 to 883 – base 760 to 3369

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttttaaggatggtatttatttttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtatatctgcggagattctactgaatgtgctaatttgcttctccaatatgg
cagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaag
tcaaacaaatgtacaaaaccccaactttgaaatattttggcggttttaatttttcacaaatattacctgaccctctaaagccaacta
agaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgcctag

gtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatgatt
gctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaatacctttg
ctatgcaataa

- MCS...6xHIS base 3370 to 3442

tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattga

- AXO1 Terminater base 3443-3783

gttttagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggat
gtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcatttt
gtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattcga
gtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

Figure 30

PGAPZ alpha 883m clone G3-7
Deduced Amino Acid Sequence (SEQ ID NO:19)

Alpha Signal Sequence -
mrfpsiftavlfaassalaapvntttedetaqipaeavigysdlegdfdvavlpfsnstnngllfinttiasiaakeegvslek
reaea

- Spike amino acids 14 to 883, does not include the first 13 amino acid leader
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfcttgigyqpyr
vvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaervdtsyecdipigagicas
yhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvdcnmyicgdstecanlllqygsf
ctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnfsqilpdplkptkrsfiedllfnkvtladagfmkqygec
lgdinardlicaqkfngltvlppllddmiaaytaalvsgtatagwtfgagaalqipfamq*

Figure 31

pGAPZ alpha 883m clone G3-7
Linear Map

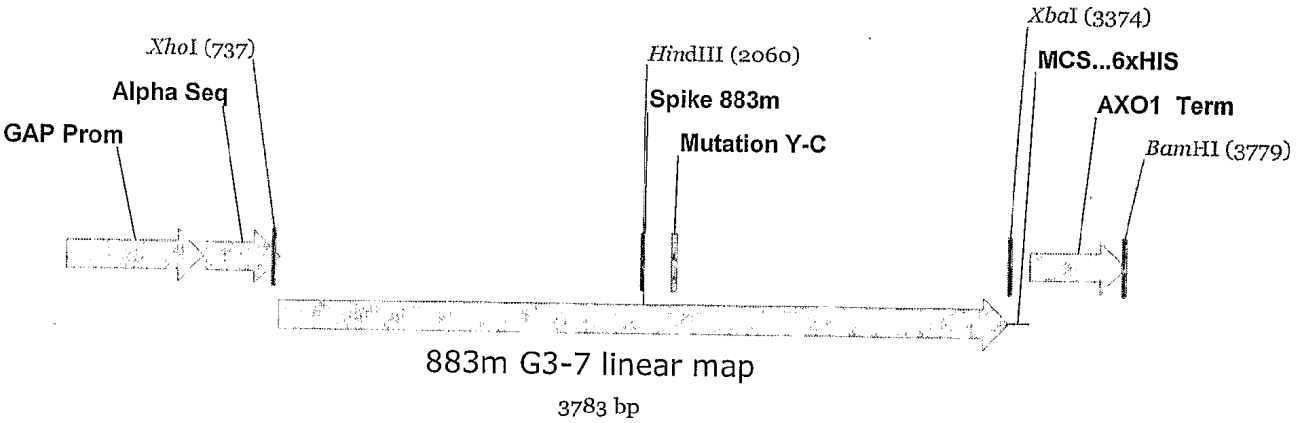


Figure 32

PGAPZ alpha 883 clone G3-7
Sequence (SEQ ID NO:20)

GAP Pomoter – base 1 to 483

agatctttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctggctccgttgcaactccgaacg
acctgctggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaacgtctcttcccttctctctcct
tccaccgcccgttaccgtccctaggaaattttactctgctggagagcttcttctacggccccccttgcagcaatgctcttcccag
cattacgttgcgggtaaaacggaggtcgtgtacccgacctagcagcccagggatggaaaagtcccggccgtcgctggca
ataatagcgggcggacgcatgtcatgagattattggaaaccaccagaatcgaatataaaaggcgaacacctttcccaatttt
ggtttctcctgacccaaagactttaaatttaatttatttgtccctatttcaatcaattgaacaactat

-Spacer - base 484 to 492

ttcgaaacg

- Alpha Signal Sequence – base 493 to 759

atgagatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatga
aacggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaac
agcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaag
agaggctgaagct

- Spike – base 760 to 3369

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttaaggatggtatttatttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgccttttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttgcaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccg
gccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactgg
tgtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatcct
aaaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagt
tgctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatatt
ctactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattc
ctattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtct
ttaggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtatatctgcggagattctactgaatgtgctaatttgcttctccaatatgg
cagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaag
tcaaacaaatgtacaaaaccccaactttgaaatattttggcggttttaatttttcacaaatattacctgaccctctaaagccaacta
agaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgcctag

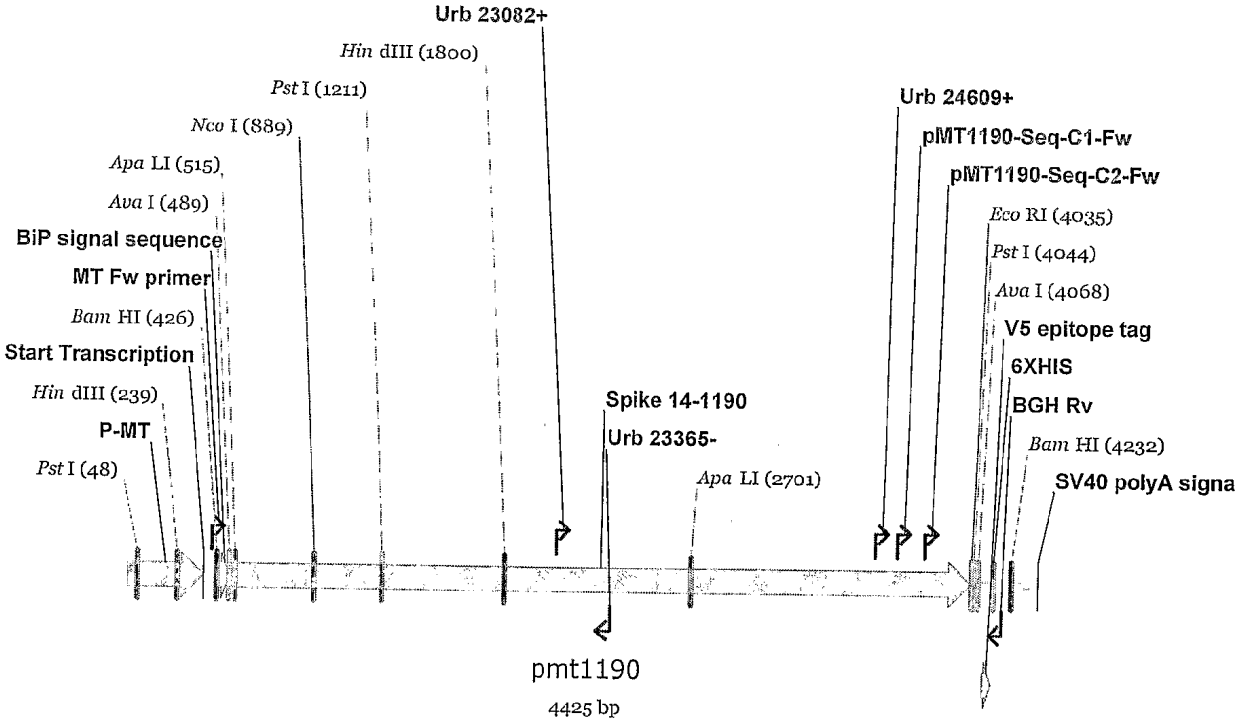
gtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatgatt
gctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaatacctttg
ctatgcaataa

- MCS...6xHIS base 3370 to 3442
tctagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattga

- AXO1 Terminater base 3443-3783
gttttagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggat
gtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcatttt
gtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattcga
gtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcc

Figure 33

pMT-Spike 1190



(SEQ ID NO:21)

P-MT Promoter

Start: 1 End: 367

gttgcaggacaggatgtggtgcccgatgtgactagctctttgctgcaggccgtcctatcctctggttccgataagagaccca
gaactccggccccccaccgcccaccgccacccccatacatatgtggtacgcaagtaagagtgcctgcgcatgccccatgt
gccccaccaagagttttgcatcccatacaagtccccaaagtggagaaccgaaccaattcttcgcgggcagaacaaaagctt
ctgcacacgtctccactcgaatttggagccggccggcgtgtgcaaaagaggtgaatcgaacgaaagacccgtgtgtaaag
ccgcgtttccaaaatgtataaaaccgagagcatctggccaatgt

BiP signal sequence

Start: 440 End: 493

atgaagttatgcatattactggccgtcgtggcctttgttggcctctcgctcggg

Spike 14-1190
Start: 500 End: 4033

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttaaggatggtatttattttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaatttttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctctttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtacatctgcggagattctactgaatgtgctaatttgcttctccaatatg
gtagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaa
gtcaaacaaatgtacaaaaccccaactttgaaatattttggtggttttaatttttcacaaatattacctgaccctctaaagccaact
aagaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgccta
ggtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatga
ttgctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaataccttt
gctatgcaaatggcatataggttcaatggcattggagttacccaaaatgttctctatgagaaccaaaaacaaatcgccaacca
atttaacaaggcgattagtcaaattcaagaatcacttacaacaacatcaactgcattgggcaagctgcaagacgttgttaacc
agaatgctcaagcattaaacacacttgttaaacaacttagctctaattttggtgcaatttcaagtgtgctaaatgatatcctttcgc
gacttgataaagtcgaggcggaggtacaaattgacaggttaattacaggcagacttcaaagccttcaaacctatgtaacaca
acaactaatcagggctgctgaaatcagggcttctgctaatcttgctgctactaaaatgtctgagtgtgttcttggacaatcaaaa
agagttgacttttgtggaaagggctaccaccttatgtccttcccacaagcagccccgcatggtgttgtcttcctacatgtcacg
tatgtgccatcccaggagaggaacttcaccacagcgccagcaatttgtcatgaaggcaaagcatacttccctcgtgaaggt
gtttttgtgtttaatggcacttcttggtttattacacagaggaacttcttttctccacaaataattactacagacaatacatttgtctc
aggaaattgtgatgtcgttattggcatcattaacaacacagtttatgatcctctgcaacctgagctcgactcattcaaagaaga
gctggacaagtacttcaaaaatcatacatcaccagatgttgatcttggcgacatttcaggcattaacgcttctgtcgtcaacatt
caaaaagaaattgaccgcctcaatgaggtcgctaaaaatttaaatgaatcactcattgaccttcaagaattgggaaaatatga
gcaataa

SV40 polyA signal
Start: 4355 End: 4360
aataaa

pMT 1190 deduced Amino Acid sequence (SEQ ID NO:22)

- BiP Signal Amino Acids
mklcillavvafvglslg

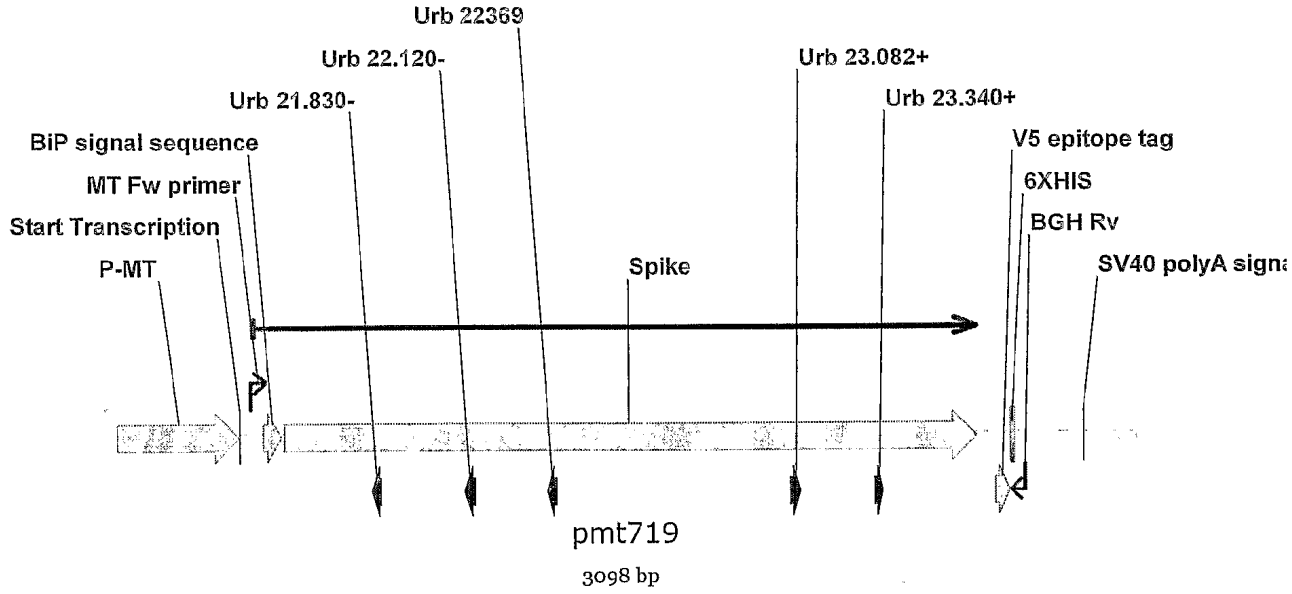
- Spike Amino Acids (14 to 1190)

sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfyttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvdcnmyicgdstecanlllqygs
fctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnfsqilpdplkptkrsfiedllfnkvtladagfmkqyge
clgdinardlicaqkfngltvlppllddmiaaytaalvsgtatagwtfgagaalqipfamqmayrfngigvtqnvlyen
qkqianqfnkaisqiqesltttstalgklqdvvnqnaqalntlvkqlssnfgaissvlndilsrldkveaevqidrlitgrlqsl
qtyvtqqliraaeirasanlaatkmsecvlgqskrvdfcgkgyhlmsfpqaaphgvvflhvtyvpsqernfttapaiche
gkayfpregvfvfngtswfitqrnffspqiittdntfvsgncdvvigiinntvydplqpeldsfkeeldkyfknhtspdvd
lgdisginasvvniqkeidrlnevaknlneslidlqelgkyeq*

Figure 34

PMT-719

pMT-719 – Map



(SEQ ID NO:23)
pMT 719 sequence, from linear map

- P-MT promoter – base 1 to 367
gttgcaggacaggatgtggtgcccgatgtgactagctctttgctgcaggccgtcctatcctctggttccgataagagaccca
gaactccggccccccaccgcccaccgccacccccatacatatgtggtacgcaagtaagagtgcctgcgcatgccccatgt
gccccaccaagagttttgcatcccatacaagtccccaaagtggagaaccgaaccaattcttcgcgggcagaacaaaagctt
ctgcacacgtctccactcgaatttggagccggccggcgtgtgcaaaagaggtgaatcgaacgaaagacccgtgtgtaaag
ccgcgtttccaaaatgtataaaaccgagagcatctggccaatgt

- BiP Signal sequence – base 440 to 493
atgaagttatgcatattactggccgtcgtggcctttgttggcctctcgctcggg

- Spike 719 – base 500 to 2617
agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttaaggatggtatttatttttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtcctttggagaggtttttaatgctactaaatt

ccettctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtataa

- SV 40 Poly A signal Transcription terminator – base 2939 to 2944
aataaa

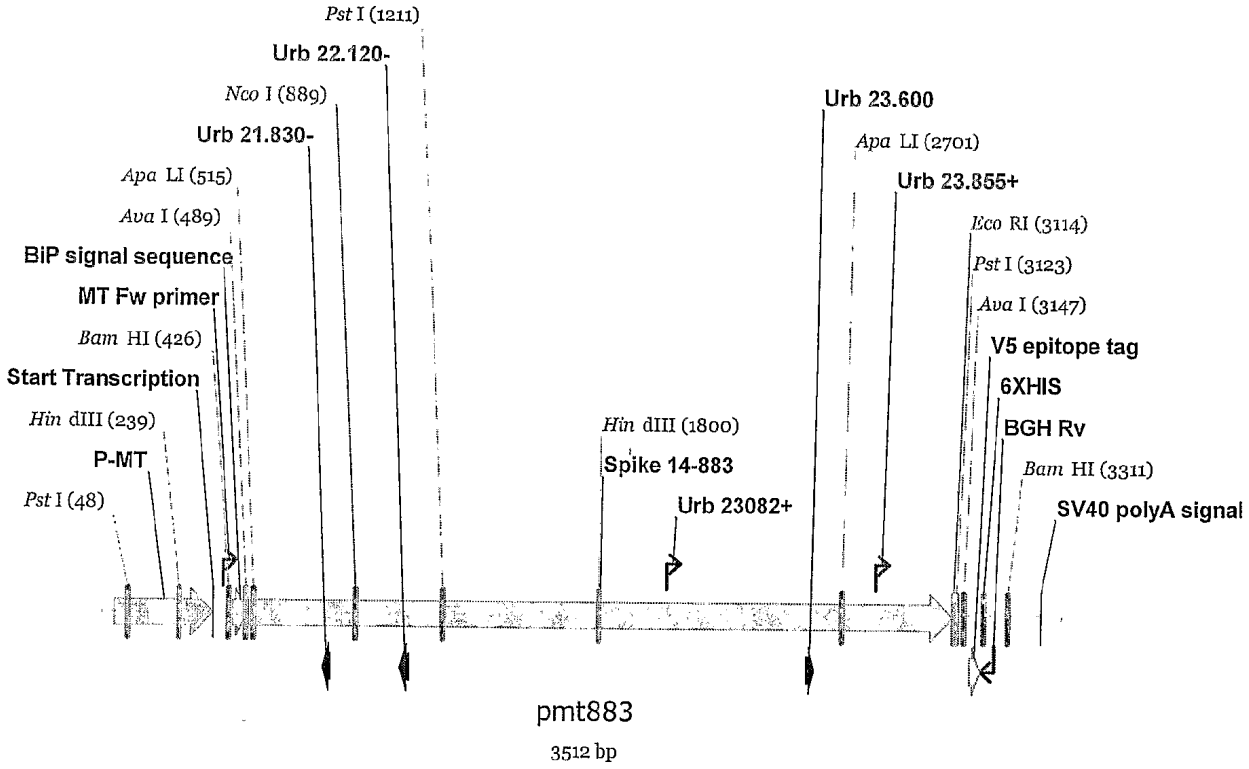
pMT 719 deduced Amino Acid sequence (SEQ ID NO:24)

- BiP Signal Amino Acids
mklcillavvafvglslg

- Spike Amino Acids (14 to 719)
sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfytttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsv*

Figure 35

PMT-Spike 883



(SEQ ID NO:25)

P-MT Promoter

Start: 1 End: 367

gttgcaggacaggatgtggtgcccgatgtgactagctctttgctgcaggccgtcctatcctctggttccgataagagaccca
gaactccggcccccccaccgcccaccgccacccccatacatatgtggtacgcaagtaagagtgcctgcgcatgccccatgt
gccccaccaagagttttgcatcccatacaagtccccaaagtggagaaccgaaccaattcttcgcgggcagaacaaaagctt
ctgcacacgtctccactcgaatttggagccggccggcgtgtgcaaaagaggtgaatcgaacgaaagacccgtgtgtaaag
ccgcgtttccaaaatgtataaaaccgagagcatctggccaatgt

BiP signal sequence

Start: 440 End: 493

atgaagttatgcatattactggccgtcgtggcctttgttggcctctcgctcggg

Spike 14-883

Start: 500 End: 3112

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcataccttttaaggatggtatttattttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtcctttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtacatctgcggagattctactgaatgtgctaatttgcttctccaatatg
gtagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaa
gtcaaacaaatgtacaaaaccccaactttgaaatattttggtggttttaatttttcacaaatattacctgaccctctaaagccaact
aagaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgccta
ggtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatga
ttgctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaataccttttt
gctatgcaataa

SV40 polyA signal
Start: 3434 End: 3439

aataaa

pMT 883 deduced Amino Acid sequence (SEQ ID NO:26)

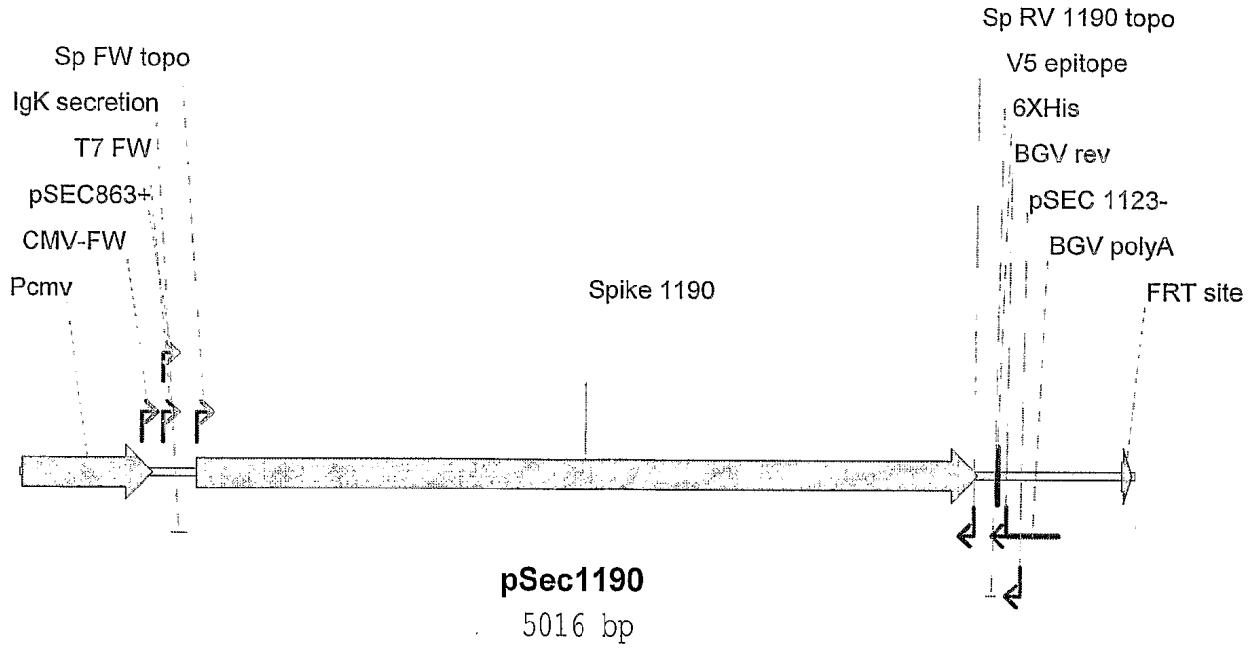
- BiP Signal Amino Acids

mklcillavvafvglslg

- **Spike Amino Acids (14 to 883)**

Sdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlflpfysnvtgfhtinhtfgnpvipfkdgiyfaateks
nvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffavskpmgtqthtmifdnafnctfeyisdafsldvse
ksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpifklplginitnfrailtafspaqdiwgtsaaayfvgy
lkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqtsnfrvvpsgdvvrfpnitnlcpfgevfnatkfpsv
yawerkkisncvadysvlynstffstfkcygvsatklndlcfsnvyadsfvvkgddvrqiapgqtgviadynyklpddf
mgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpfspdgkpctppalncywplndygfytttgigyqpy
rvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltpsskrfqpfqqfgrdvsdftdsvrdpktseildispcsf
ggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriystgnnvfqtqagcligaehvdtsyecdipigagica
syhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisittevmpvsmaktsvdcnmyicgdstecanlllqygs
fctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnfsqilpdplkptkrsfiedllfnkvtladagfmkqyge
clgdinardlicaqkfngltvlppllddmiaaytaalvsgtatagwtfgagaalqipfamq*

Figure 36



(SEQ ID NO:27)

Pcmv

Start: 1 End: 588

gttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttccgcgttacat
aacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagt
aacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatc
atatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatggga
ctttcctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggat
agcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggac
tttccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcaga
gctc

IgK secretion

atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggtgac

Start: 674 End: 736

Spike 1190

Start: 782 End: 4312

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttaaggatggtatttatttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg

tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtacatctgcggagattctactgaatgtgctaatttgcttctccaatatg
gtagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaa
gtcaaacaaatgtacaaaaccccaactttgaaatattttggtggttttaatttttcacaaatattacctgaccctctaaagccaact
aagaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgccta
ggtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatga
ttgctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaatacctttt
gctatgcaaatggcatataggttcaatggcattggagttacccaaaatgttctctatgagaaccaaaaacaaatcgccaacca
atttaacaaggcgattagtcaaattcaagaatcacttacaacaacatcaactgcattgggcaagctgcaagacgttgttaacc
agaatgctcaagcattaaacacacttgttaaacaacttagctctaattttggtgcaatttcaagtgtgctaaatgatatcctttcgc
gacttgataaagtcgaggcggaggtacaaattgacaggttaattacaggcagacttcaaagccttcaaacctatgtaacaca
acaactaatcagggctgctgaaatcagggcttctgctaatcttgctgctactaaaatgtctgagtgtgttcttggacaatcaaaa
agagttgacttttgtggaaagggctaccaccttatgtccttcccacaagcagccccgcatggtgttgtcttcctacatgtcacg
tatgtgccatcccaggagaggaacttcaccacagcgccagcaatttgtcatgaaggcaaagcatacttccctcgtgaaggt
gtttttgtgtttaatggcacttcttggtttattacacagaggaacttcttttctccacaaataattactacagacaatacatttgtctc
aggaaattgtgatgtcgttattggcatcattaacaacacagtttatgatcctctgcaacctgagctcgactcattcaaagaaga
gctggacaagtacttcaaaaatcatacatcaccagatgttgatcttggcgacatttcaggcattaacgcttctgtcgtcaacatt
caaaaagaaattgaccgcctcaatgaggtcgctaaaaatttaaatgaatcactcattgaccttcaagaattgggaaaatatga
gcaa

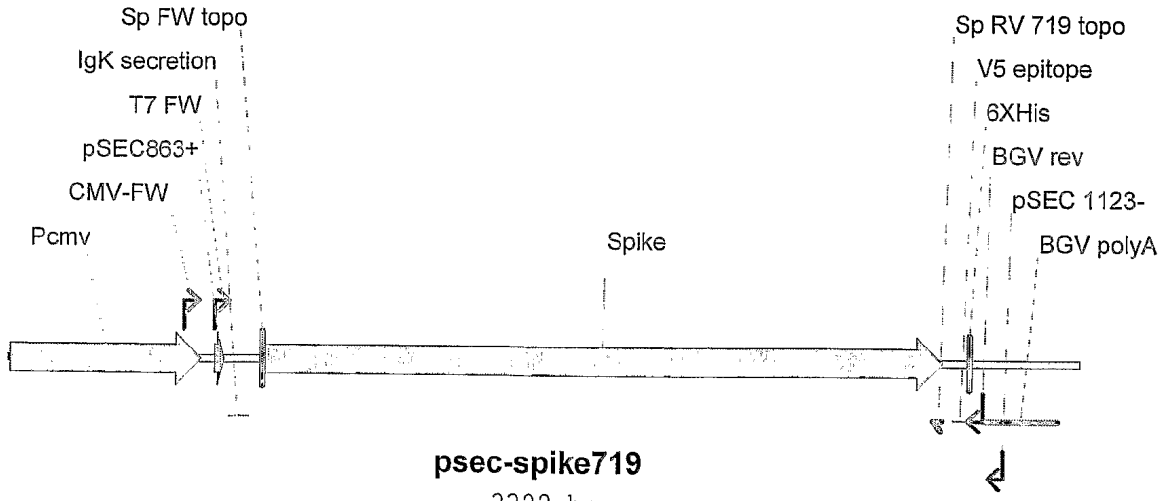
V5 epitope
Start: 4349 End: 4390
Ggtaagcctatccctaaccctctcctcggtctcgattctacg

6XHis
Start: 4400 End: 4417
catcatcaccatcaccat

BGV polyA
Start: 4446 End: 4670
ctgtgccttctagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtcct
ttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagca
agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatgg

Translation of pSec1190 (signal peptide is underlined) **(SEQ ID NO:28)**
metdtlllwvlllwvpgstgdaaqparrarrtklalsdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlfl
pfysnvtgfhtinhtfgnpvipfkdgiyfaateksnvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffav
skpmgtqthtmifdnafnctfeyisdafsldvseksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpif
klplginitnfrailtafspaqdiwgtsaaayfvgylkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqt
snfrvvpsgdvvrfpnitnlcpfgevfnatkfpsvyawerkkisncvadysvlynstffstfkcygvsatklndlcfsnv
yadsfvvkgddvrqiapgqtgviadynyklpddfmgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpf
spdgkpctppalncywplndygfytttgigyqpyrvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltps
skrfqpfqqfgrdvsdftdsvrdpktseildispcsfggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriy
stgnnvfqtqagcligaehvdtsyecdipigagicasyhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisitt
evmpvsmaktsvdcnmyicgdstecanlllqygsfctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnf
sqilpdplkptkrsfiedllfnkvtladagfmkqygeclgdinardlicaqkfngltvlppllтddmiaaytaalvsgtatag
wtfgagaalqipfamqmayrfngigvtqnvlyenqkqianqfnkaisqiqesltttstalgklqdvvnqnaqalntlvk
qlssnfgaissvlndilsrldkveaevqidrlitgrlqslqtyvtqqliraaeirasanlaatkmsecvlgqskrvdfcgkgy
hlmsfpqaaphgvvflhvtyvpsqernfttapaichegkayfpregvfvfngtswfitqrnffspqiittdntfvsgncdv
vigiinntvydplqpeldsfkeeldkyfknhtspdvdlgdisginasvvniqkeidrlnevaknlneslidlqelgkyeq

Figure 37



(SEQ ID NO:29)

Promoter cmv

Start: 1 End: 588

gttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttccgcgttacat
aacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagt
aacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatc
atatgccaagtacgcccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatggga
ctttcctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggat
agcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggac
tttccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcaga
gctc

IgK secretion

atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggtgac

Start: 674 End: 736

Spike

Start: 782 End: 2899

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttttaaggatggtatttatttttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctatttttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt
tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc

tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaactttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtataa

V5 epitope
Start: 2933 End: 2974
ggtaagcctatccctaaccctctcctcggtctcgattctacg

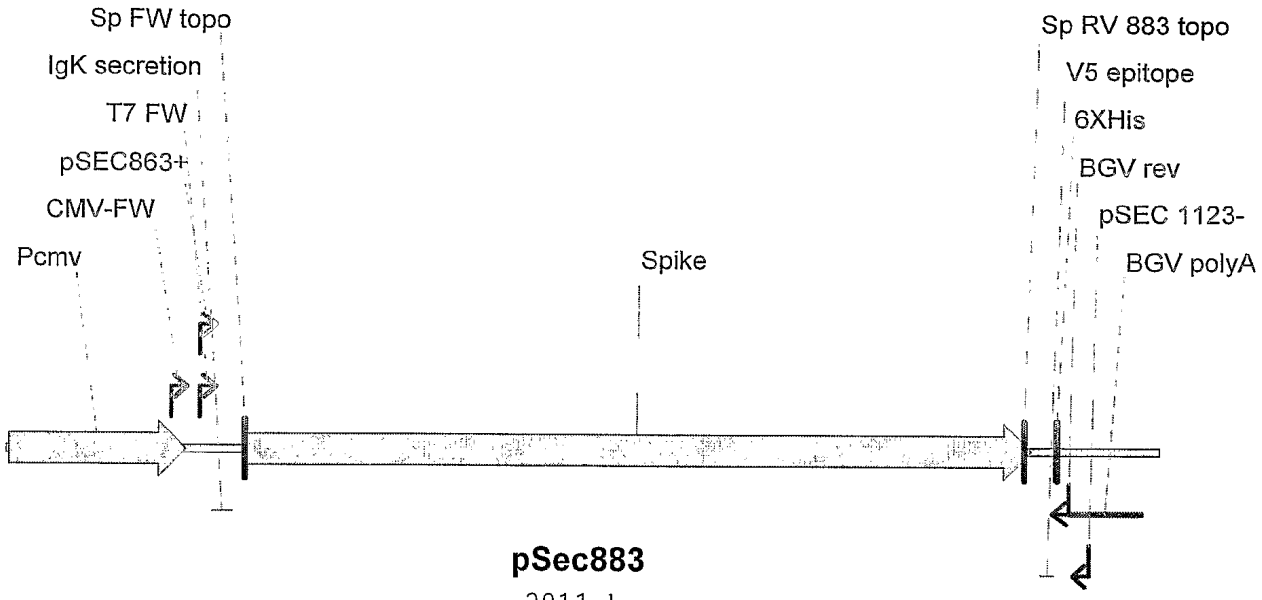
6XHis
Start: 2984 End: 3001
catcatcaccatcaccat

BGV polyA
Start: 3030 End: 3254
ctgtgccttctagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtcct
ttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagca
agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatgg

Translation of psec-spike719 (signal peptide is underlined) **(SEQ ID NO:30)**

Metdtlllwvlllwvpgstgdaaqparrarrtklalsdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlfl
pfysnvtgfhtinhtfgnpvipfkdgiyfaateksnvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffav
skpmgtqthtmifdnafnctfeyisdafsldvseksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpif
klplginitnfrailtafspaqdiwgtsaaayfvgylkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqt
snfrvvpsgdvvrfpnitnlcpfgevfnatkfpsvyawerkkisncvadysvlynstffstfkcygvsatklndlcfsnv
yadsfvvkgddvrqiapgqtgviadynyklpddfmgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpf
spdgkpctppalncywplndygfyttigigyqpyrvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltps
skrfqpfqqfgrdvsdftdsvrdpktseildispcsfggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriy
stgnnvfqtqagcligaehvdtsyecdipigagicasyhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisitt
evmpvsmaktsv

Figure 38



(SEQ ID NO:31)

Pcmv

Start: 1 End: 588

gttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttccgcgttacat
aacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagt
aacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatc
atatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatggga
ctttcctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggat
agcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggac
tttccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcaga
gctc

IgK secretion

atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggtgac

Start: 674 End: 736

Spike

Start: 782 End: 3394

agtgaccttgaccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttact
atcctgatgaaatttttagatcagacactctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactatt
aatcatacgtttggcaaccctgtcatacctttaaggatggtatttatttgctgccacagagaaatcaaatgttgtccgtggttg
ggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgttatacgagcatgtaacttt
gaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcatttaattgc
actttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaa
ataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacc
tatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacg
tcagctgcagcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgt

tgattgttctcaaaatccacttgctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaattt
cagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtccttttggagaggtttttaatgctactaaatt
cccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaactcaacatttttttcaacc
tttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaagggagatg
atgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtcctt
gcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggc
cctttgagagagacatatctaatgtgccttttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccatta
aatgattatggtttttacaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccgg
ccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggt
gtgttaactccttcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatttcactgattccgttcgagatccta
aaacatctgaaatattagacatttcaccttgcgcttttgggggtgtaagtgtaattacacctggaacaaatgcttcatctgaagtt
gctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattc
tactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacattcc
tattggagctggcatttgtgctagttaccatacagtttctttattacgtagtactagccaaaaatctattgtggcttatactatgtctt
taggtgctgatagttcaattgcttactctaataacaccattgctatacctactaacttttcaattagcattactacagaagtaatgc
ctgtttctatggctaaaacctccgtagattgtaatatgtacatctgcggagattctactgaatgtgctaatttgcttctccaatatg
gtagcttttgcacacaactaaatcgtgcactctcaggtattgctgctgaacaggatcgcaacacacgtgaagtgttcgctcaa
gtcaaacaaatgtacaaaaccccaactttgaaatattttggtggttttaatttttcacaaatattacctgaccctctaaagccaact
aagaggtcttttattgaggacttgctctttaataaggtgacactcgctgatgctggcttcatgaagcaatatggcgaatgccta
ggtgatattaatgctagagatctcatttgtgcgcagaagttcaatggacttacagtgttgccacctctgctcactgatgatatga
ttgctgcctacactgctgctctagttagtggtactgccactgctggatggacatttggtgctggcgctgctcttcaaataccttttt
gctatgcaataa

V5 epitope
Start: 3428 End: 3469
ggtaagcctatccctaaccctctcctcggtctcgattctacg

6XHis
Start: 3479 End: 3496
catcatcaccatcaccat

BGV polyA
Start: 3525 End: 3749
ctgtgccttctagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtcct
ttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagca
agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatgg

Translation of psec-spike883 (signal peptide is underlined) **(SEQ ID NO:32)**
metdtlllwvlllwvpgstgdaaqparrarrtklalsdldrcttfddvqapnytqhtssmrgvyypdeifrsdtlyltqdlfl
pfysnvtgfhtinhtfgnpvipfkdgiyfaateksnvvrgwvfgstmnnksqsviiinnstnvviracnfelcdnpffav
skpmgtqthtmifdnafnctfeyisdafsldvseksgnfkhlrefvfknkdgflyvykgyqpidvvrdlpsgfntlkpif
klplginitnfrailtafspaqdiwgtsaaayfvgylkpttfmlkydengtitdavdcsqnplaelkcsvksfeidkgiyqt
snfrvvpsgdvvrfpnitnlcpfgevfnatkfpsvyawerkkisncvadysvlynstffstfkcygvsatklndlcfsnv
yadsfvvkgddvrqiapgqtgviadynyklpddfmgcvlawntrnidatstgnynykyrylrhgklrpferdisnvpf
spdgkpctppalncywplndygfytttgigyqpyrvvvlsfellnapatvcgpklstdliknqcvnfnfngltgtgvltps
skrfqpfqqfgrdvsdftdsvrdpktseildispcafggvsvitpgtnassevavlyqdvnctdvstaihadqltpawriy
stgnnvfqtqagcligaehvdtsyecdipigagicasyhtvsllrstsqksivaytmslgadssiaysnntiaiptnfsisitt
evmpvsmaktsvdcnmyicgdstecanlllqygsfctqlnralsgiaaeqdrntrevfaqvkqmyktptlkyfggfnf

sqilpdplkptkrsfiedllfnkvtladagfmkqygeclgdinardlicaqkfngltvlpplltddmiaaytaalvsgtatag
wtfgagaalqipfamq*

Figure 39

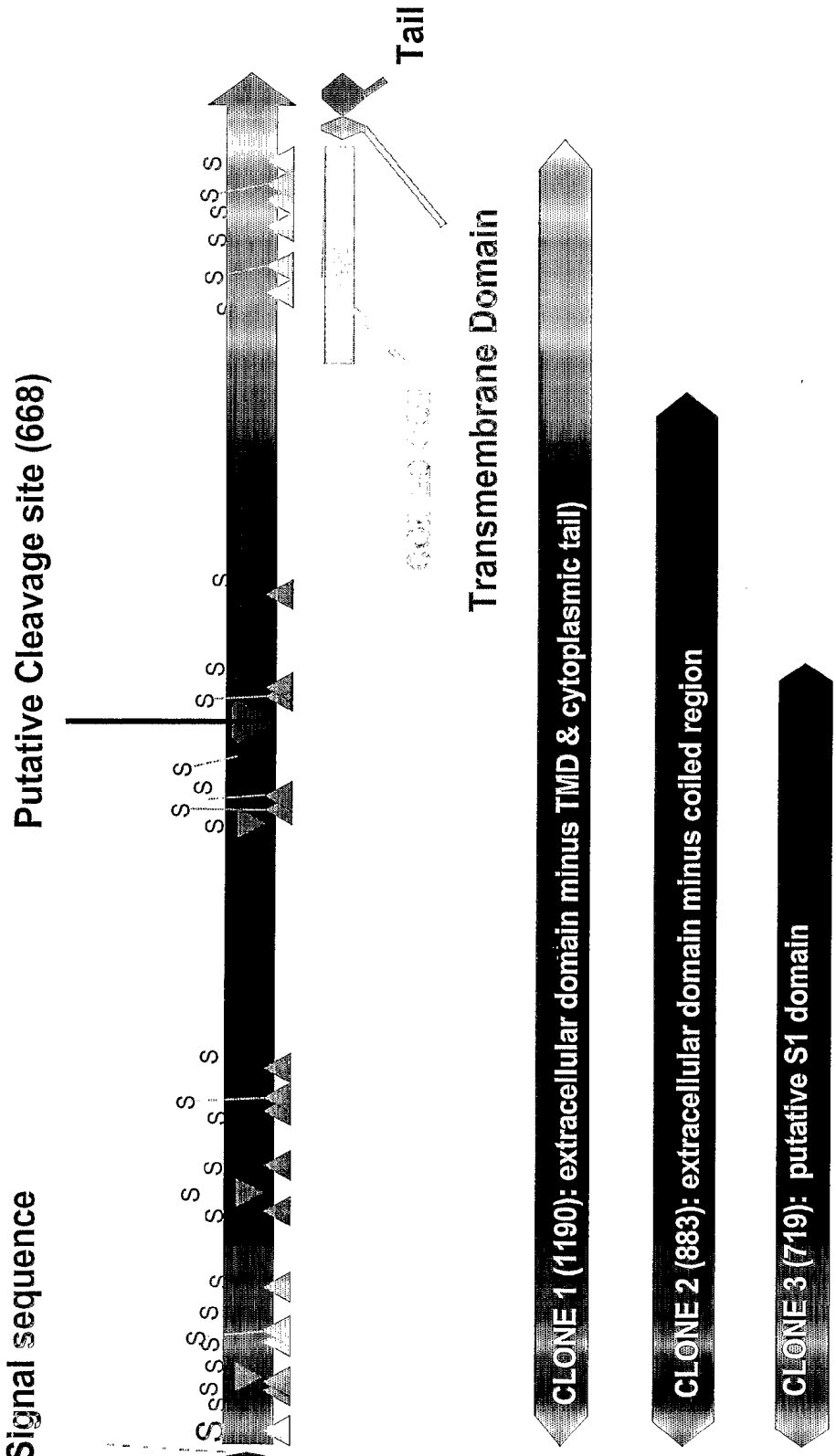

Signal sequence
Putative Cleavage site (668)
Tail
Transmembrane Domain
CLONE 1 (1190): extracellular domain minus TMD & cytoplasmic tail)
CLONE 2 (883): extracellular domain minus coiled region
CLONE 3 (719): putative S1 domain

Figure 40

Characterized expression clones
Lipofection / electroporation
Stably integrated, multiple copies on genome
Drosophila S2 Schneider
Cell suspension
Inducible
Secreted
Stably integrated, multiple copies on genome
Pichia pastoris
Cell suspension
Inducible
Constitutive
Secreted
Secreted

Figure 41

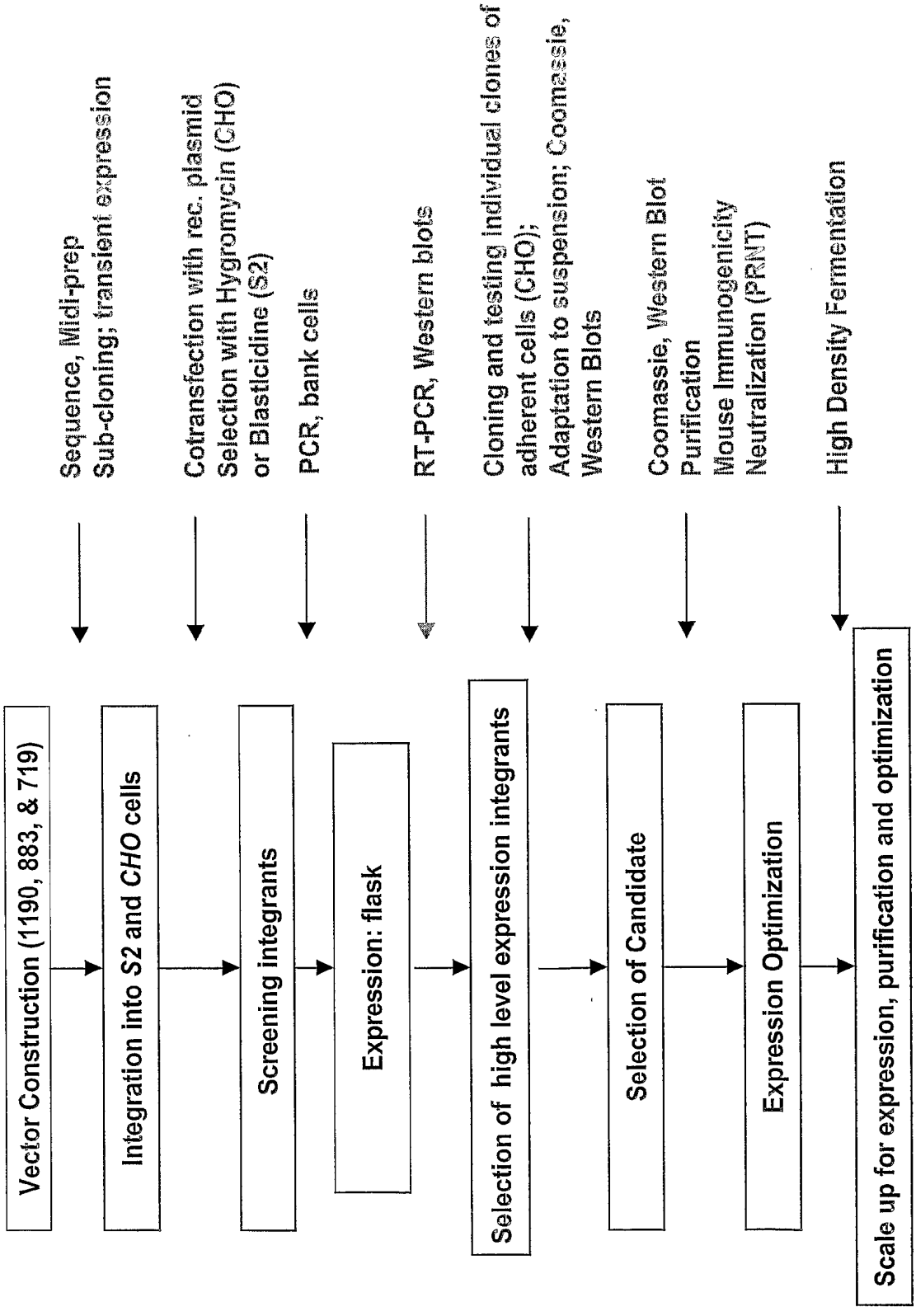
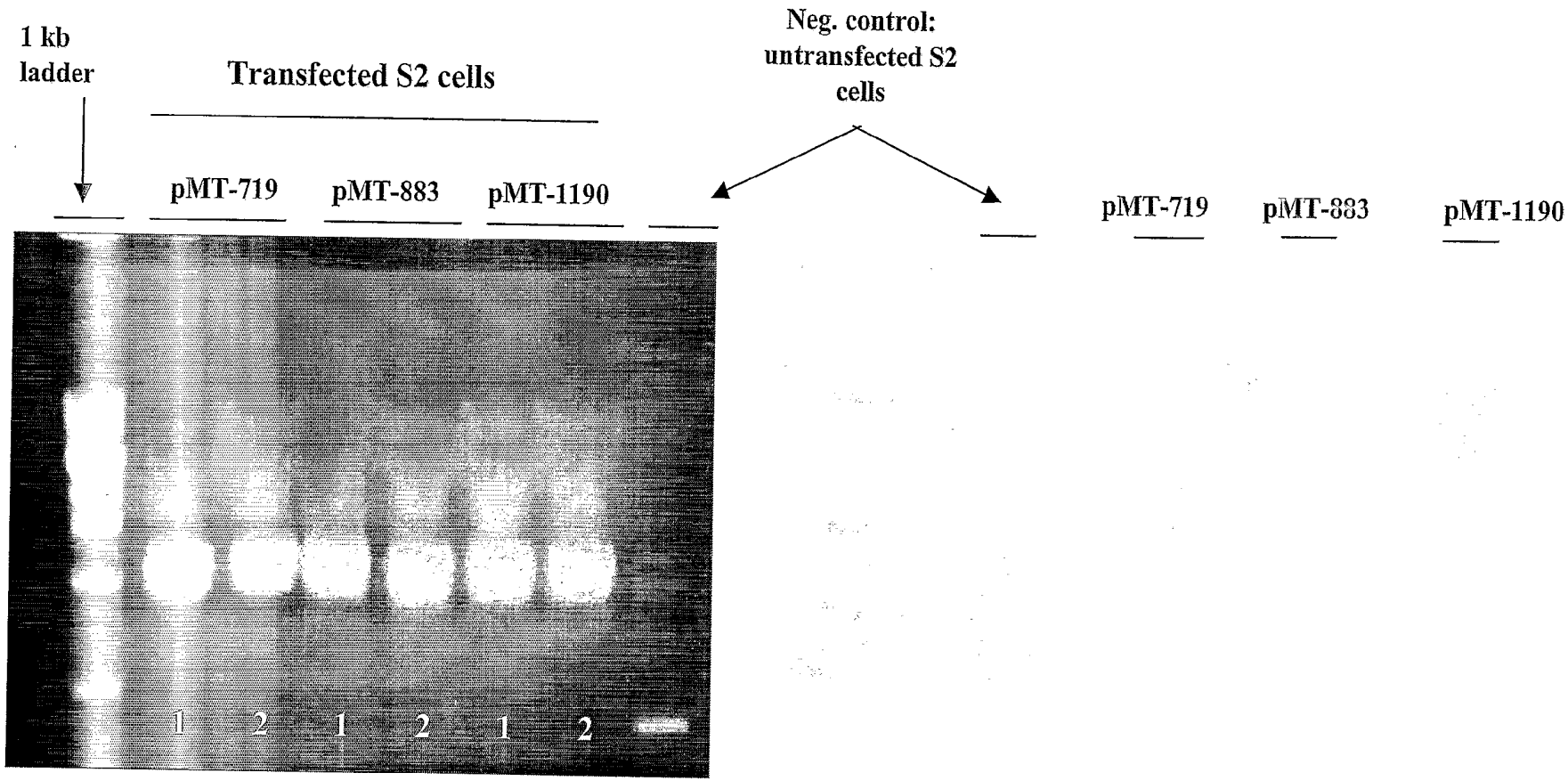

Vector Construction (1190, 883, & 719)
Integration into S2 and CHO cells
Screening integrants
Expression: flask
Selection of high level expression integrants
Selection of Candidate
Expression Optimization
Scale up for expression, purification and optimization
Sequence, Midi-prep
Sub-cloning; transient expression
Cotransfection with rec. plasmid
Selection with Hygromycin (CHO)
or Blasticidine (S2)
PCR, bank cells
RT-PCR, Western blots
Cloning and testing individual clones of
adherent cells (CHO);
Adaptation to suspension; Coomassie,
Western Blots
Coomassie, Western Blot
Purification
Mouse Immunogenicity
Neutralization (PRNT)
High Density Fermentation

Figure 42



For each construct:

Lane 1: 24h after transfection

Lane 2: 48h after transfection

Figure 43

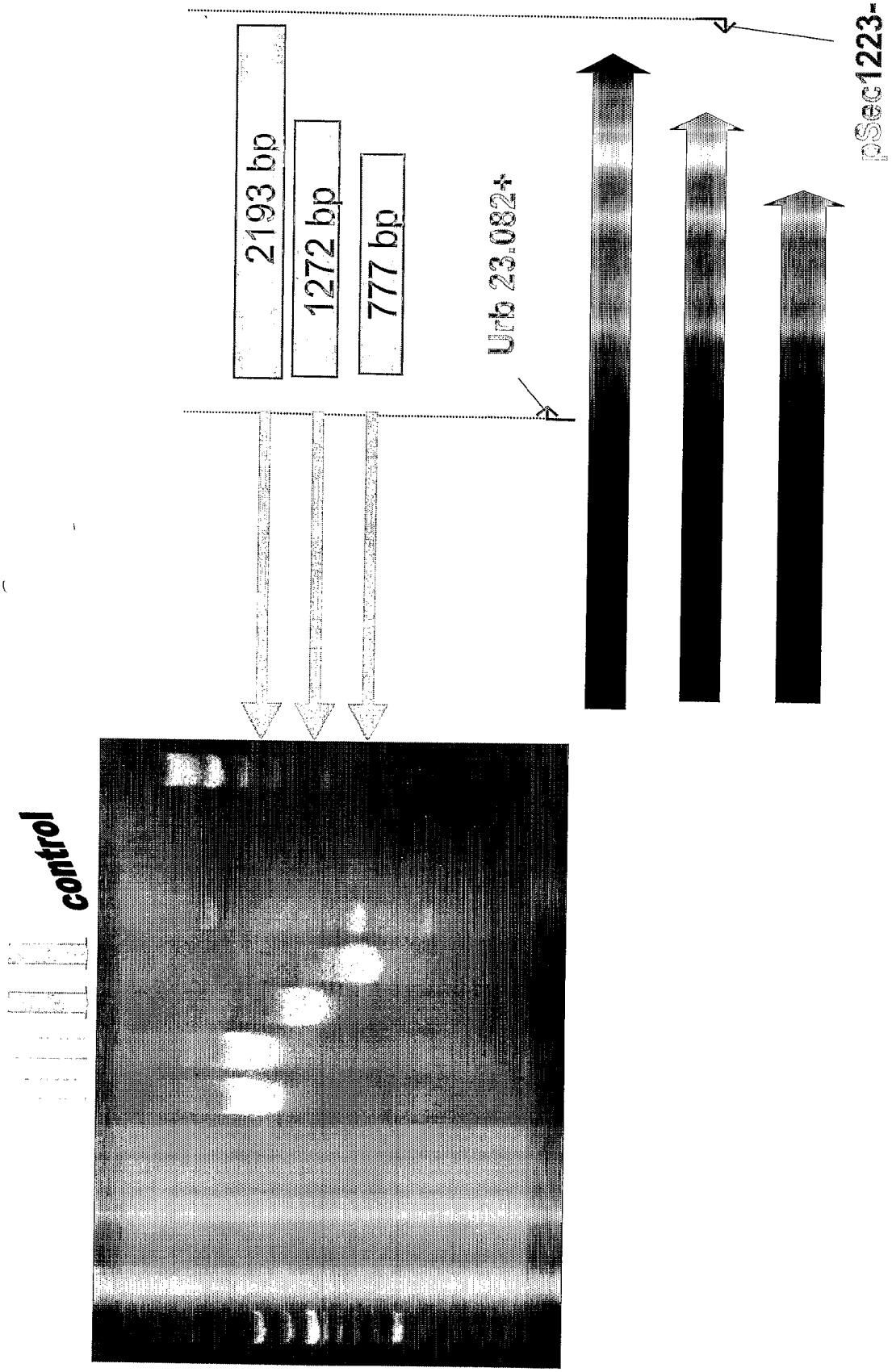


Figure 44

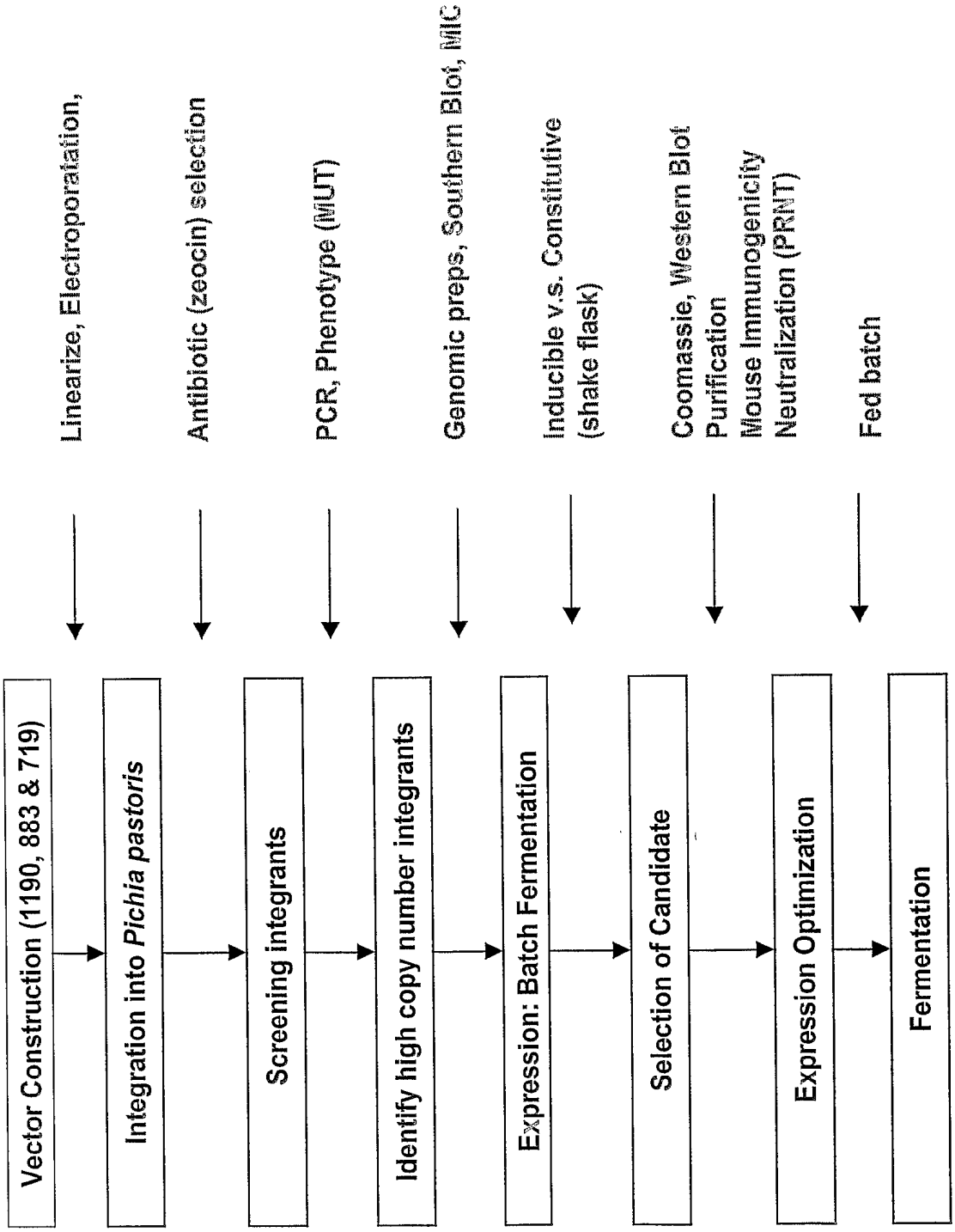

Vector Construction (1190, 883 & 719)
Linearize, Electroporation,
Integration into Pichia pastoris
Antibiotic (zeocin) selection
Screening integrants
PCR, Phenotype (MUT)
Identify high copy number integrants
Genomic preps, Southern Blot, MIC
Expression: Batch Fermentation
Inducible v.s. Constitutive (shake flask)
Selection of Candidate
Coomassie, Western Blot
Purification
Mouse Immunogenicity
Neutralization (PRNT)
Expression Optimization
Fed batch
Fermentation

Figure 45 (page 1 of 2)

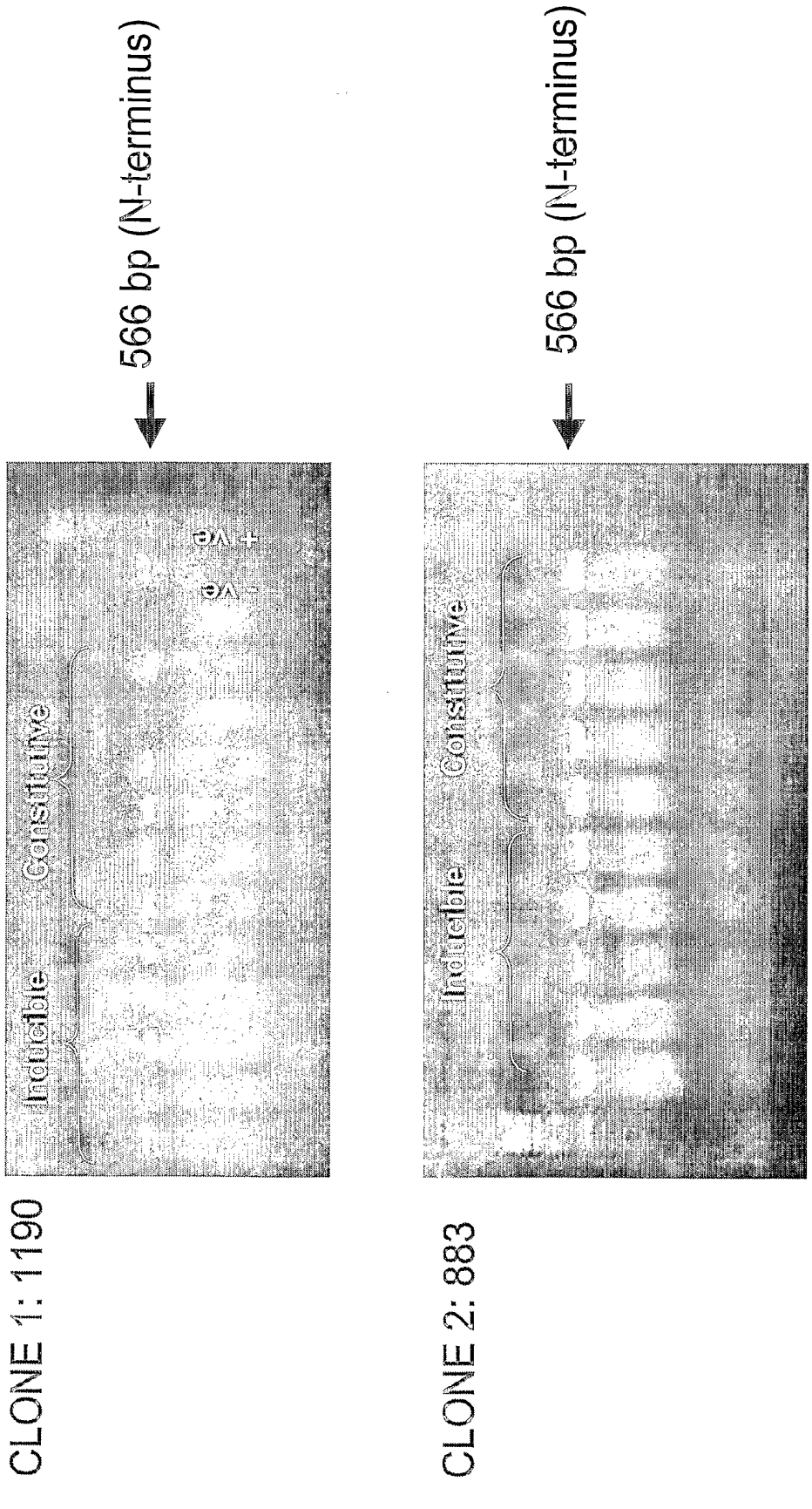


Figure 45 (page 2 of 2)

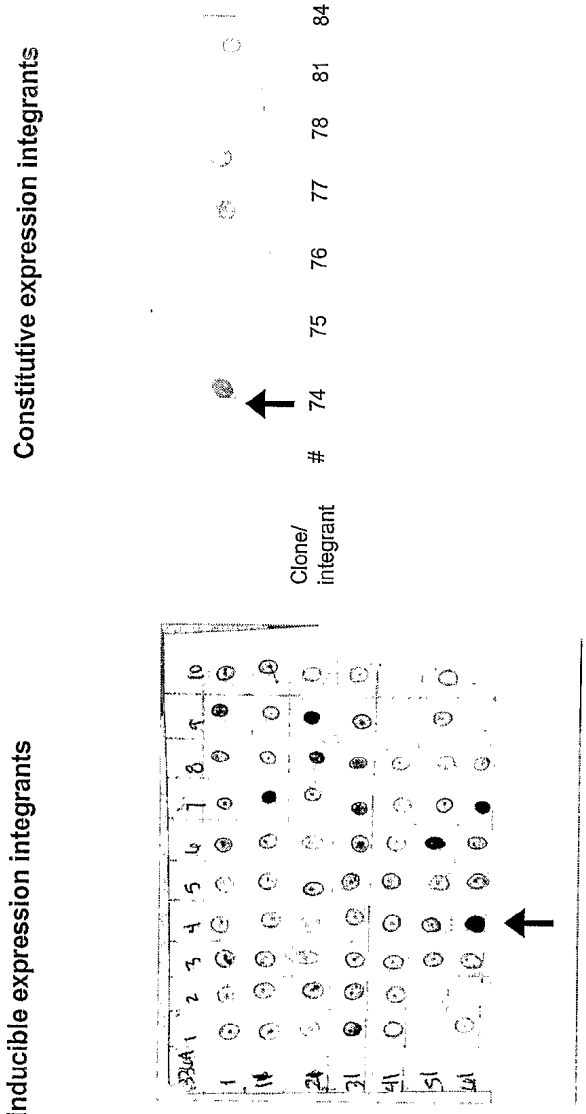


Figure 46 (page 1 of 2)

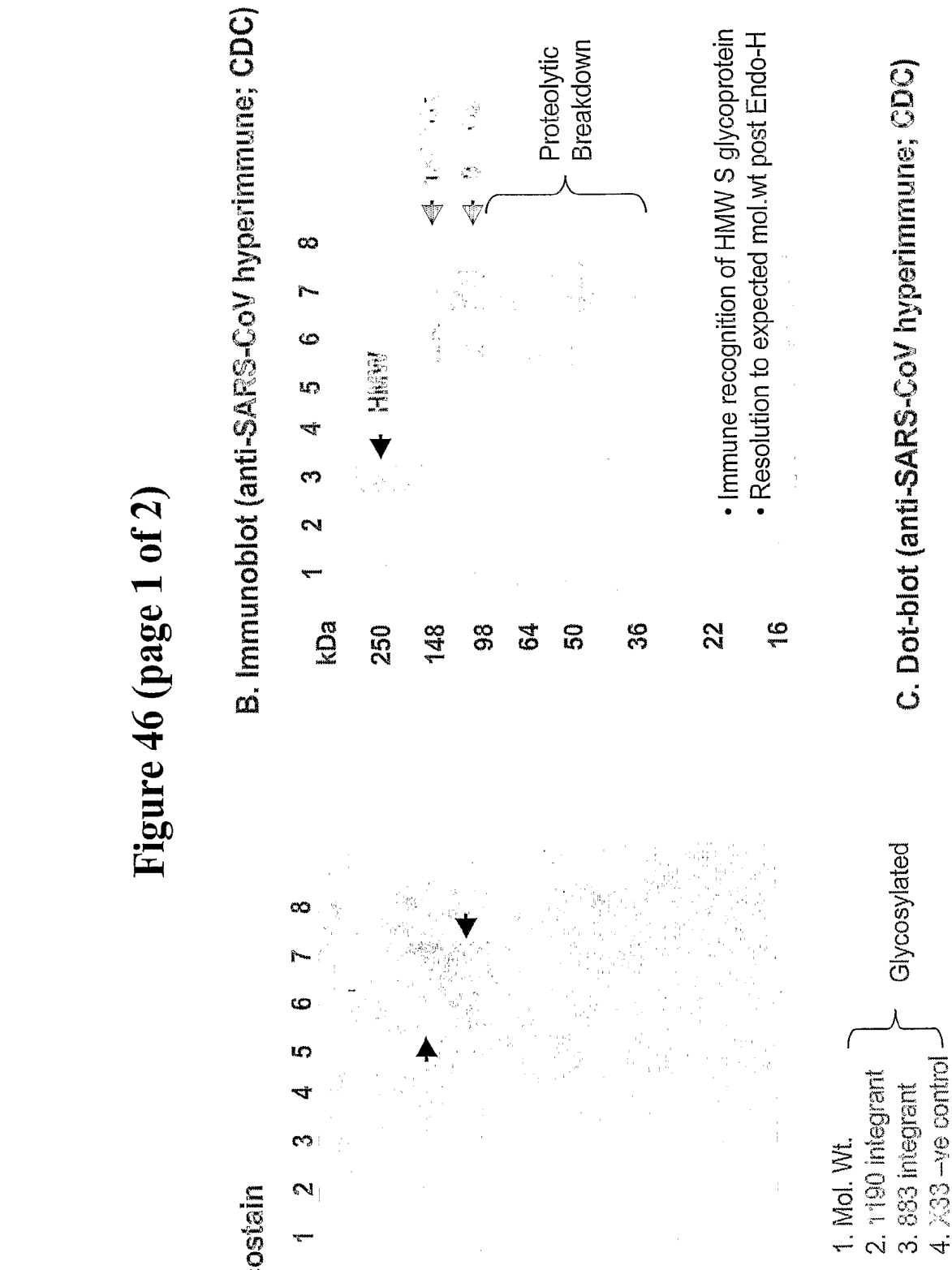
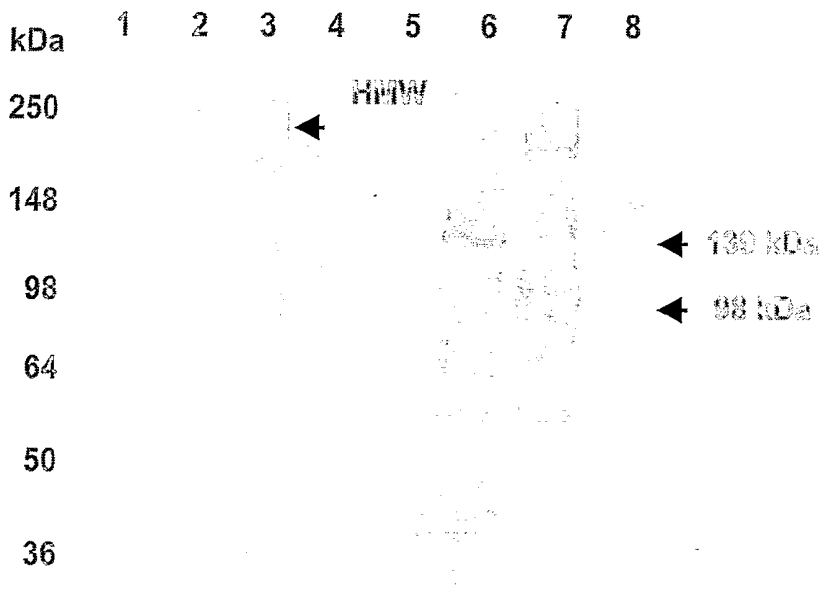
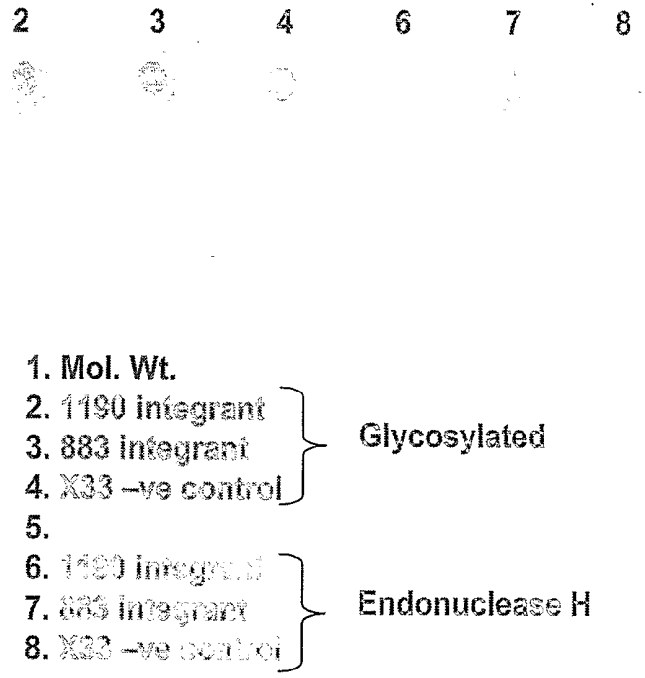


Figure 46 (page 2 of 2)

D. Immunoblot (human convalecent sera)

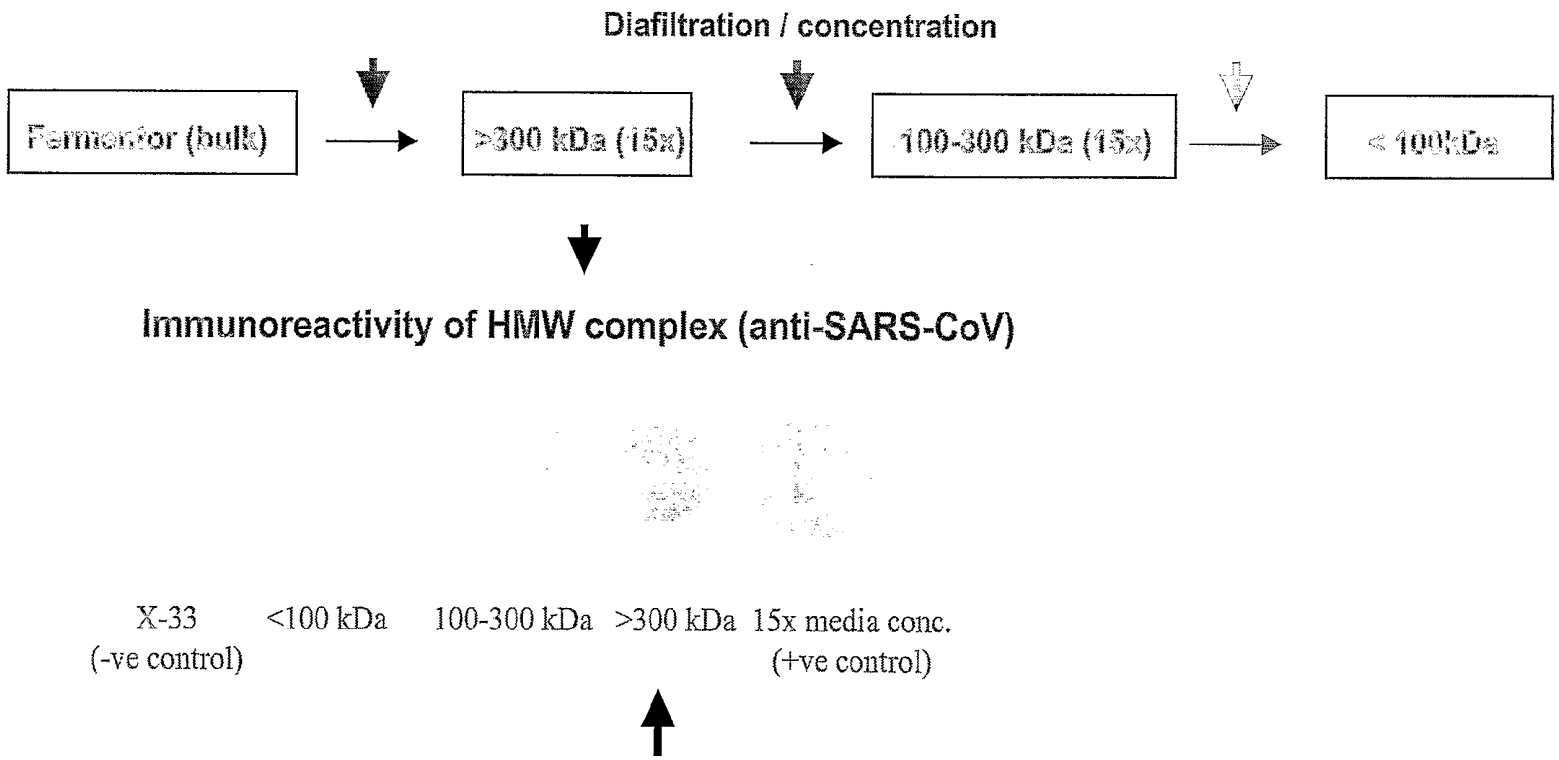


E. Dot-blot (human convalecent sera)



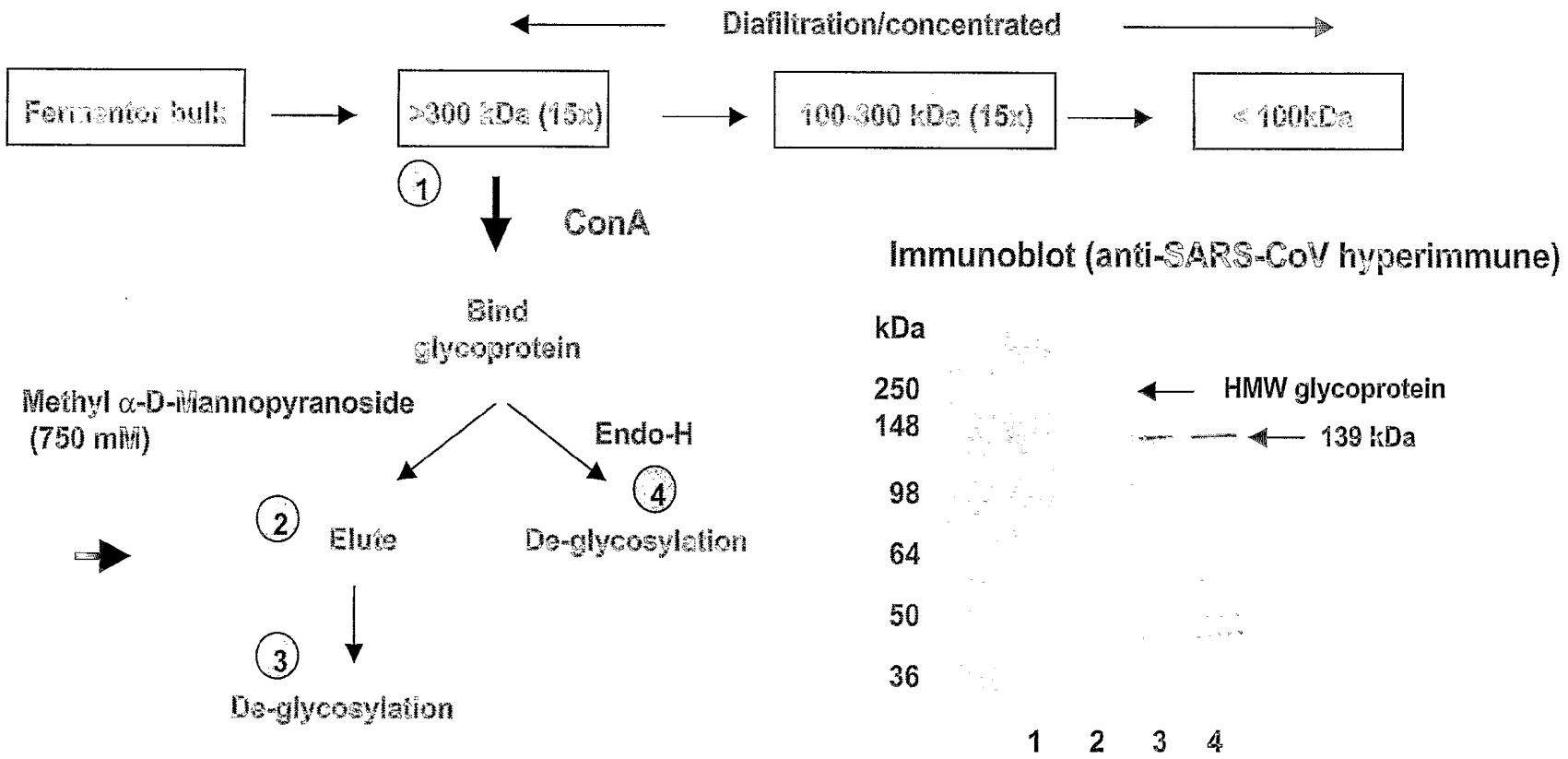
- Human immune sera recognizes HMW & de-glycosylated 1190 & 883 S glycoprotein (expected mol. wt)
- Evidence for enhanced recognition of conformational epitopes

Figure 47



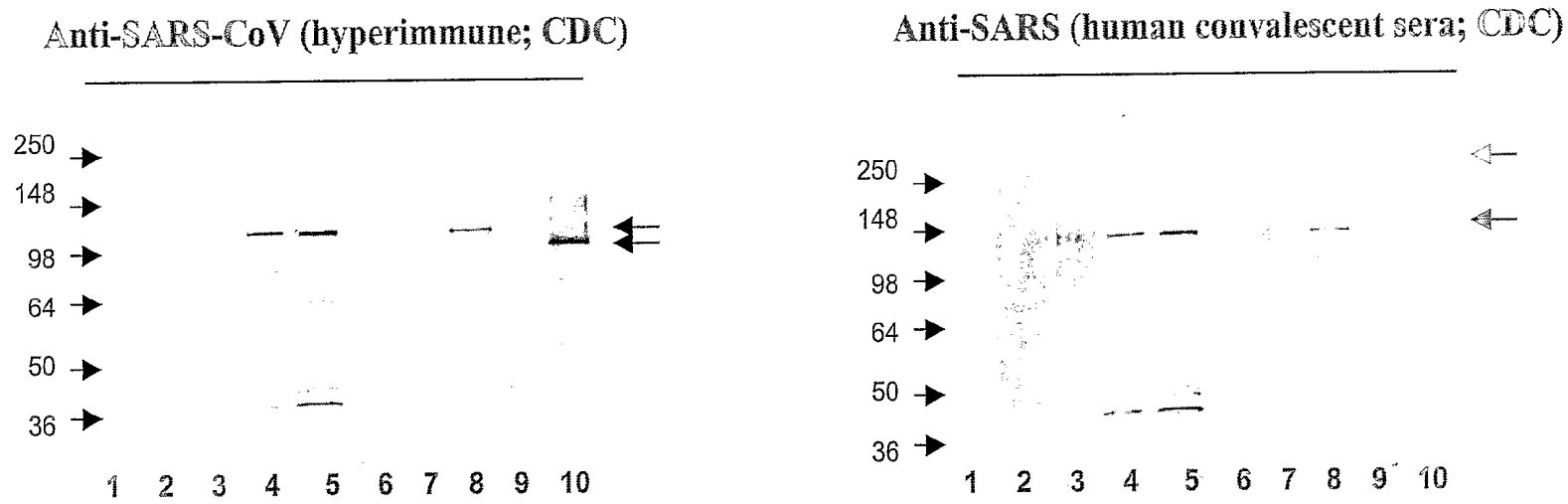
- recombinant S protein > 300 kDa
- Native HMW S glycoprotein immunoreactive with anti-SARS-CoV hyperimmune (CDC)

Figure 48



- recognition of HMW complex by Western blot at > 250 kDa with anti-SARS-CoV (CDC)
- Lectin binding of recombinant S glycoprotein
- recombinant S glycoprotein resolves to expected mol. wt post endonuclease treatment (139 kDa)

Figure 49



1. Marker
2. pGAP-1190 : >300kD fraction
3. : >300 kD fraction , Con A elution
4. : >300 kD fraction , Con A elution + Endo-H treatment
5. : >300 kD fraction , Endo-H Elution

6. X33 negative control (*Pichia pastoris*)

7. Q Seph. (+DTT)
8. Q Seph. (+DTT + Endo-H)

9. Q Seph. (-DTT)
10. Q Seph. (-DTT + Endo-H)

- Glycosylated monomer (-DTT, lane 9) migrates at >250 kDa

Figure 50

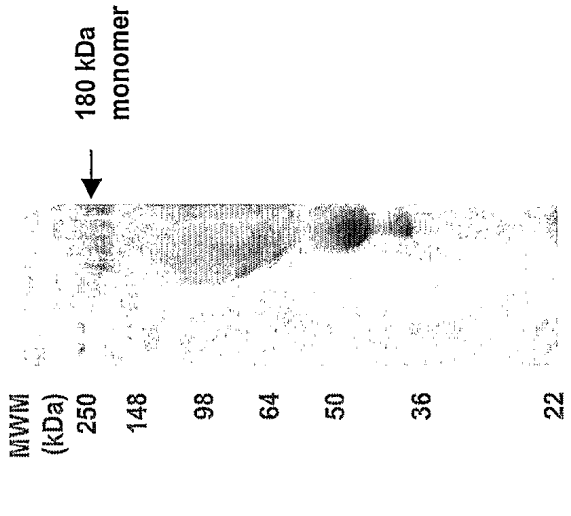
NCBInr 20031016 (1539396 sequences; 498883957 residues)
Significant hits: E2 glycoprotein precursor; putative spike glycoprotein [SARS coronavirus] ACID PHOSPHATASE PHO1 PRECURSOR Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) trypsin (EC 3.4.21.4) precursor - bovine Modified Beta Trypsin (Monoisopropylphosphoryl Inhibited) (E.C.3.4.21.4) (Neutron (SEQ ID NO:33)

Matched peptides shown in **Bold Red**

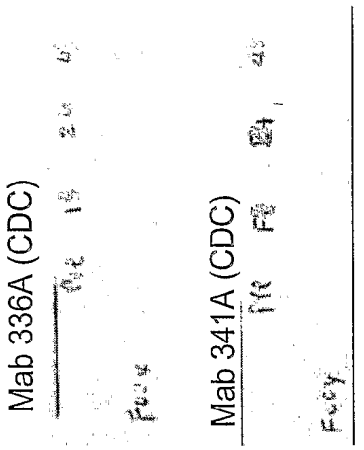
```
1    MFIFLLFLTL TSGSDLDRCT TFDDVQAPNY TQHTSSMRGV YYPDEIFRSD
51   TLYLTQDLFL PFYSNVTGFH TINHTFGNPV IPFKDGIYFA ATEKSNVVRG
101  WVFGSTMNNK SQSVIIINNS TNVVIRACNF ELCDNPFFAV SKPMGTQTHT
151  MIFDNAFNCT FEYISDAFSL DVSEKSGNFK HLREFVFKNK DGFLYVYKGY
201  QPIDVVRDLP SGFNTLKPIF KLPLGINITN FRAILTAFSP AQDIWGTSAA
251  AYFVGYLKPT TFMLKYDENG TITDAVDCSQ NPLAELKCSV KSFEIDKGIY
301  QTSNFRVVPS GDVVRFPNIT NLCPFGEVFN ATKFPSVYAW ERKKISNCVA
351  DYSVLYNSTF FSTFKCYGVS ATKLNDLCFS NVYADSFVVK GDDVRQIAPG
401  QTGVIADYNY KLPDDFMGCV LAWNTRNIDA TSTGNYNYKY RYLRHGKLRP
451  FERDISNVPF SPDGKPCTPP ALNCYWPLND YGFYTTTGIG YQPYRVVVLS
501  FELLNAPATV CGPKLSTDLI KNQCVNFNFN GLTGTGVLTP SSKRFQPFQQ
551  FGRDVSDFTD SVRDPKTSEI LDISPCAFGG VSVITPGTNA SSEVAVLYQD
601  VNCTDVSTAI HADQLTPAWR IYSTGNNVFQ TQAGCLIGAE HVDTSYECDI
651  PIGAGICASY HTVSLLRSTS QKSIVAYTMS LGADSSIAYS NNTIAIPTNF
701  SISITTEVMP VSMAKTSVDC NMYICGDSTE CANLLLQYGS FCTQLNRALS
751  GIAAEQDRNT REVFAQVKQM YKTPTLKYFG GFNFSQILPD PLKPTKRSFI
801  EDLLFNKVTL ADAGFMKQYG ECLGDINARD LICAQKFNGL TVLPPLLTDD
851  MIAAYTAALV SGTATAGWTF GAGAALQIPF AMQMAYRFNG IGVTQNVLYE
901  NQKQIANQFN KAISQIQESL TTTSTALGKL QDVVNQNAQA LNTLVKQLSS
951  NFGAISSVLN DILSRLDKVE AEVQIDRLIT GRLQSLQTYV TQQLIRAAEI
1001 RASANLAATK MSECVLGQSK RVDFCGKGYH LMSFPQAAPH GVVFLHVTYV
1051 PSQERNFTTA PAICHEGKAY FPREGVFVFN GTSWFITQRN FFSPQIITTD
1101 NTFVSGNCDV VIGIINNTVY DPLQPELDSF KEELDKYFKN HTSPDVDLGD
1151 ISGINASVVN IQKEIDRLNE VAKNLNESLI DLQELGKYEQ YIKWPWYVWL
1201 GFIAGLIAIV MVTILLCCMT SCCSCLKGAC SCGSCCKFDE DDSEPVLKGV
1251 KLHYT
```

Probability Based Mowse Score
Score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 50 indicate identity or extensive homology (p<0.05)

Figure 51


180 kDa
monomer
MWM
(kDa)
250
148
98
64
50
36
22

A.


Mab 336A (CDC)
Mab 341A (CDC)

B.

Figure 52 (page 1 of 2)

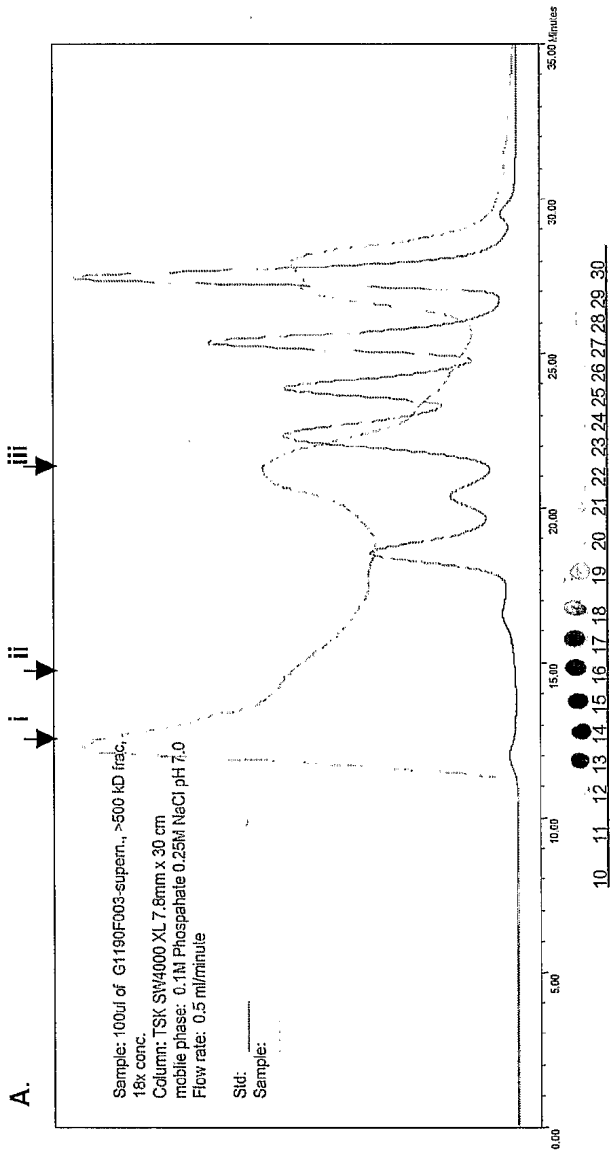


Figure 52 (page 2 of 2)

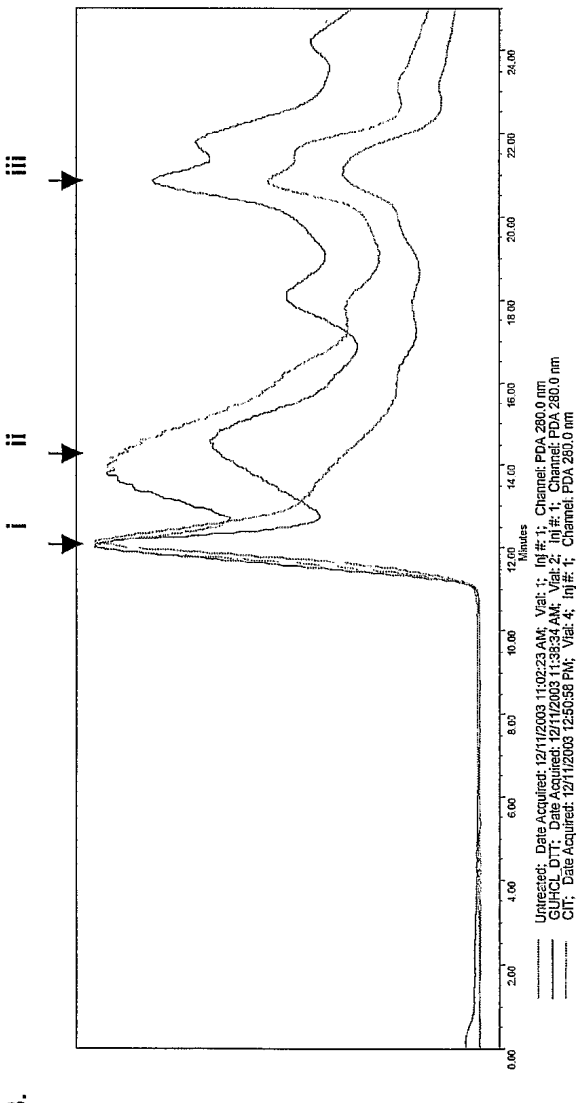
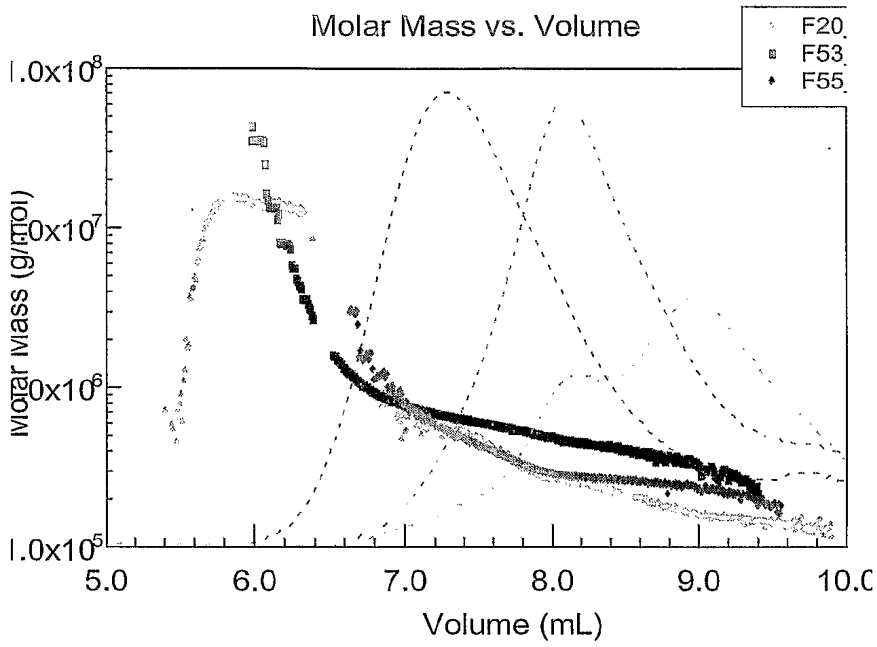


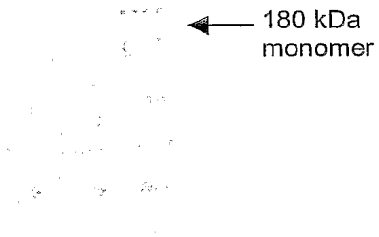
Figure 53



Fraction	MW (Average)
F 20 GuHCl/DTT	1. 322 kDa 2. 160 kDa
F 53 GuHCl/DTT to Citrate (pH4)	312 kDa
F 55 GuHCl/DTT to citrate (pH4)	623 kDa

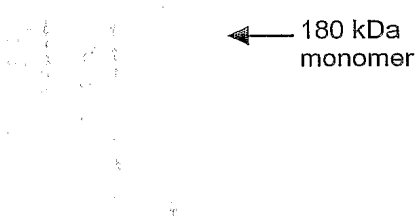
Figure 54

A. Coomassie stain (SDS-PAGE)



9 12 Marker
Fermentation day

B. Immunoblot (anti-SARS-CoV polyclonal)



9 12 Marker
Fermentation Day

Figure 55

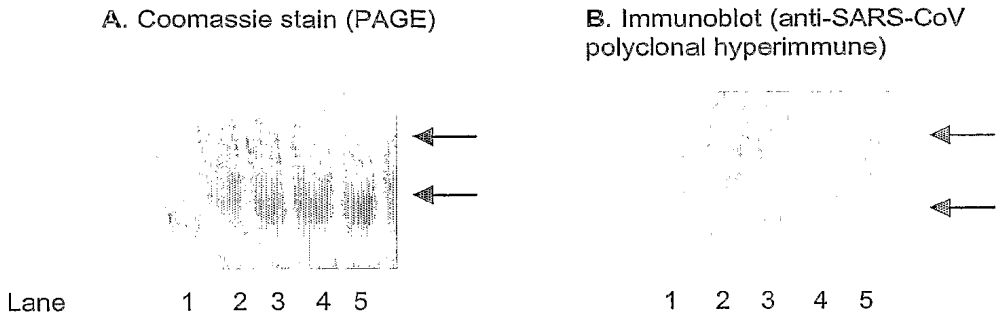

A. Coomassie stain (PAGE)
B. Immunoblot (anti-SARS-CoV polyclonal hyperimmune)
Lane
1 2 3 4 5
1 2 3 4 5

Figure 56

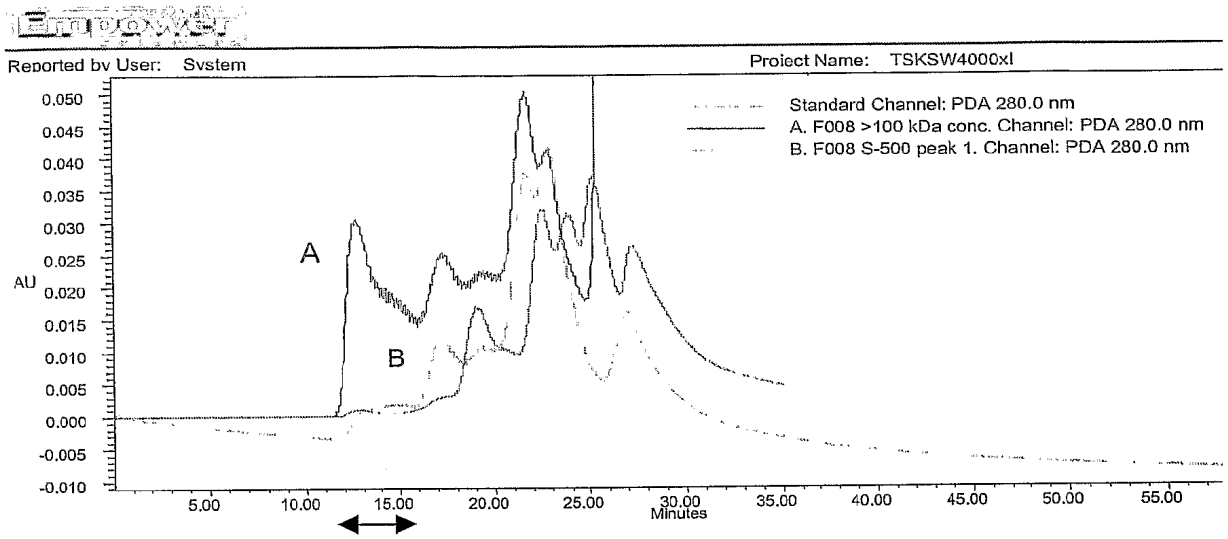

Reported by User: System
Project Name: TSKSW4000xl
Standard Channel: PDA 280.0 nm
A. F008 >100 kDa conc. Channel: PDA 280.0 nm
B. F008 S-500 peak 1. Channel: PDA 280.0 nm
AU
0.050
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
-0.005
-0.010
A
B
5.00
10.00
15.00
20.00
25.00
30.00
35.00
40.00
45.00
50.00
55.00
Minutes

Figure 57

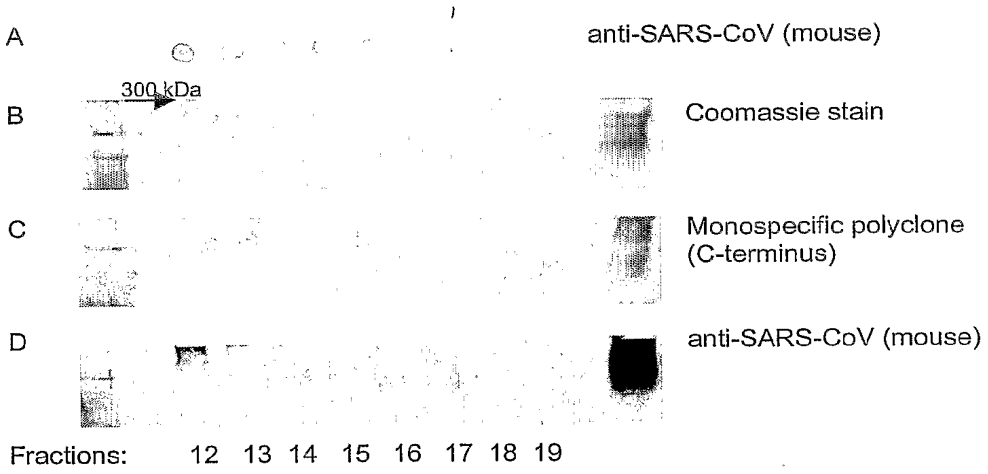
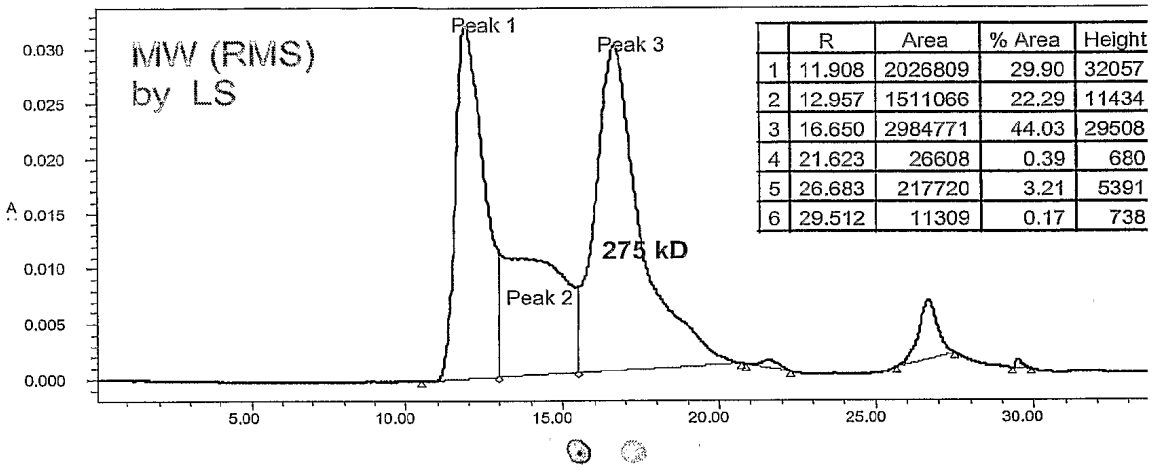

A
anti-SARS-CoV (mouse)
300 kDa
B
Coomassie stain
C
Monospecific polyclone
(C-terminus)
D
anti-SARS-CoV (mouse)
Fractions: 12 13 14 15 16 17 18 19

Figure 58



	R	Area	% Area	Height
1	11.908	2026809	29.90	32057
2	12.957	1511066	22.29	11434
3	16.650	2984771	44.03	29508
4	21.623	26608	0.39	680
5	26.683	217720	3.21	5391
6	29.512	11309	0.17	738

Figure 59

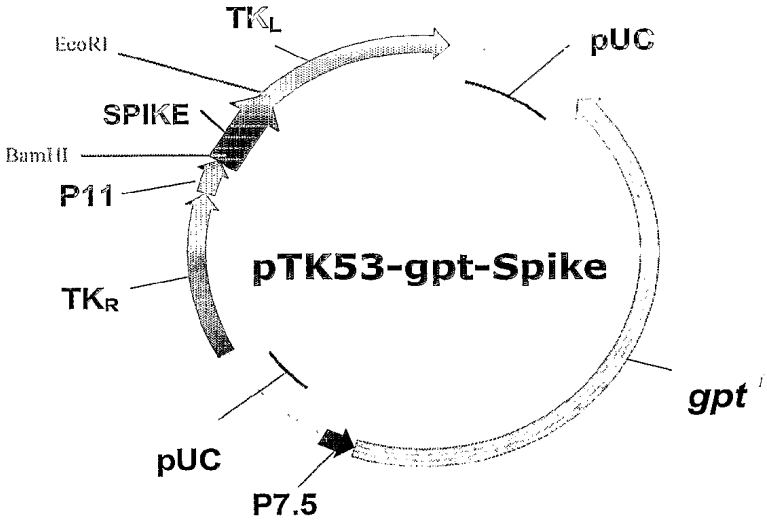


Figure 60

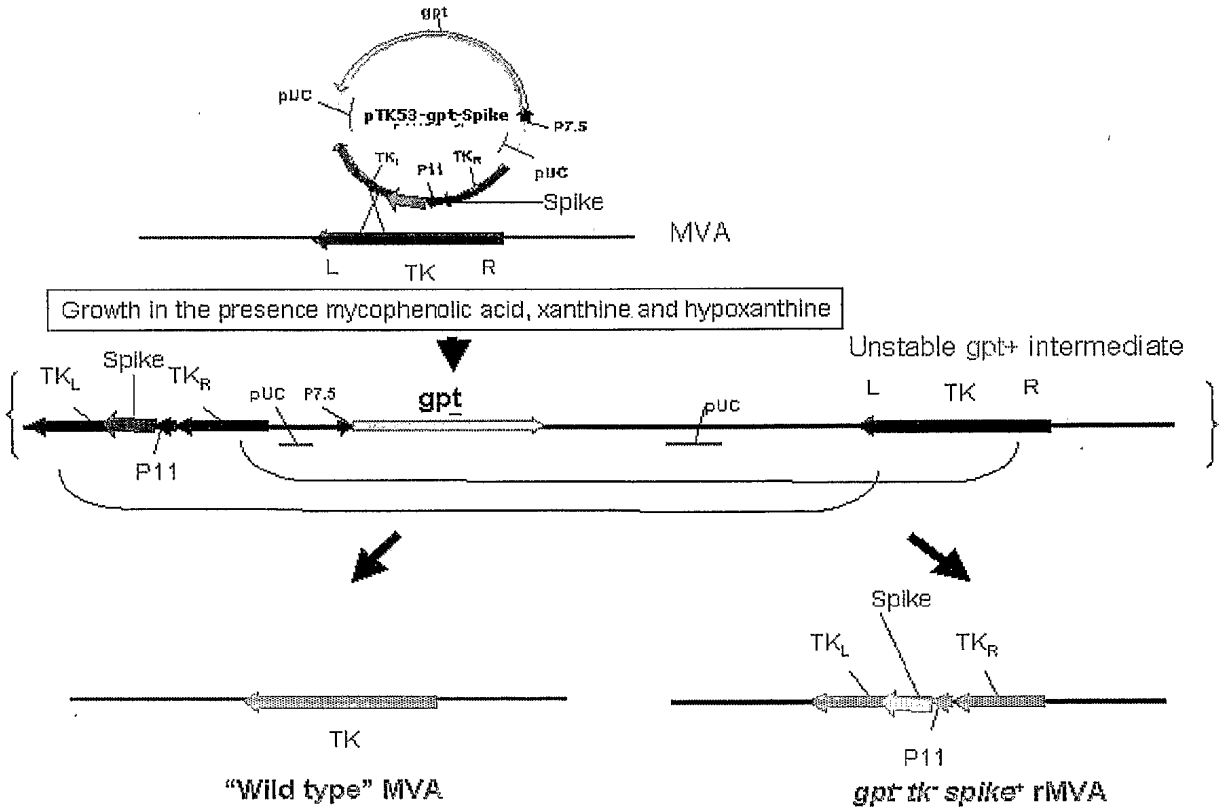


Figure 61

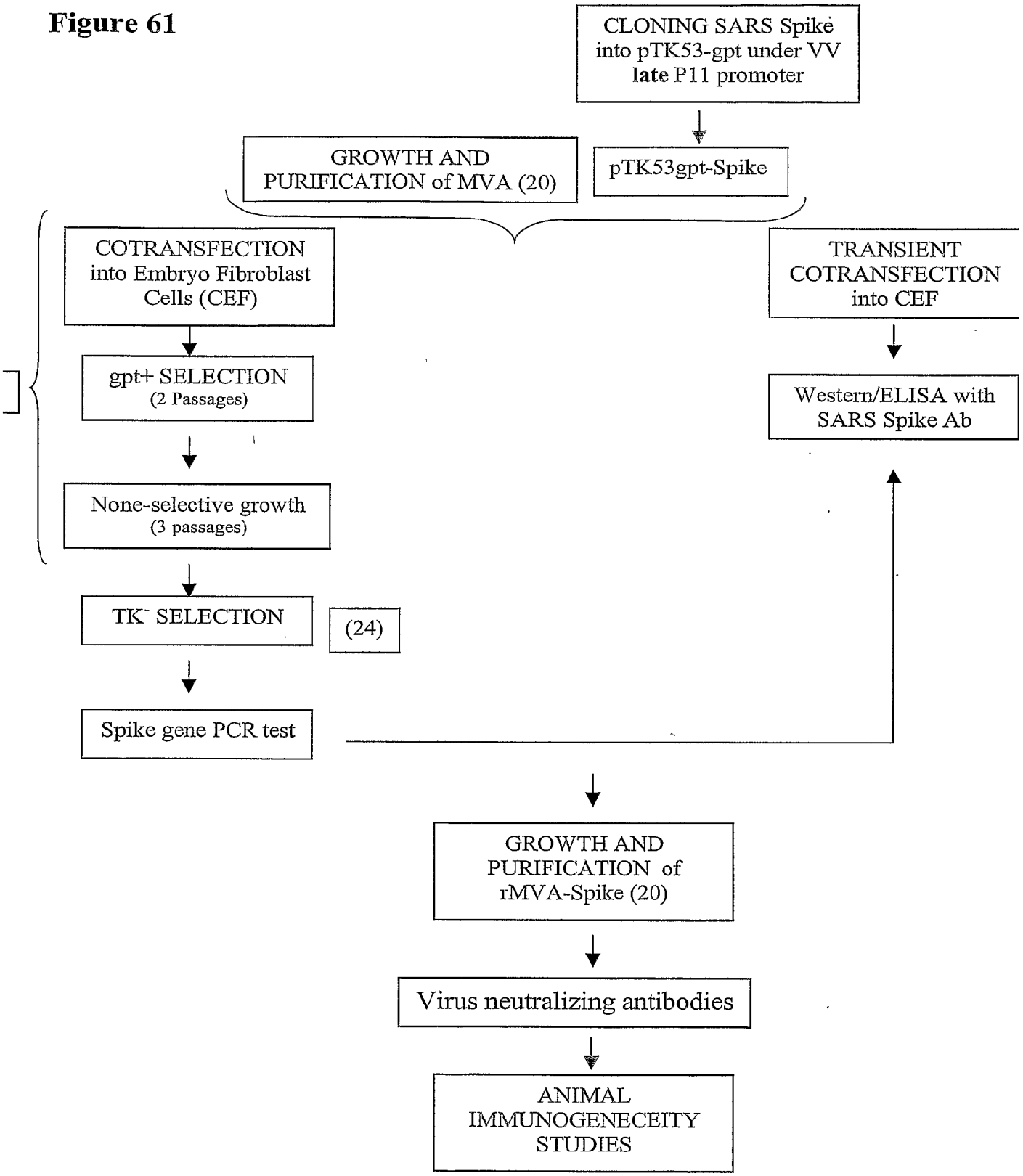
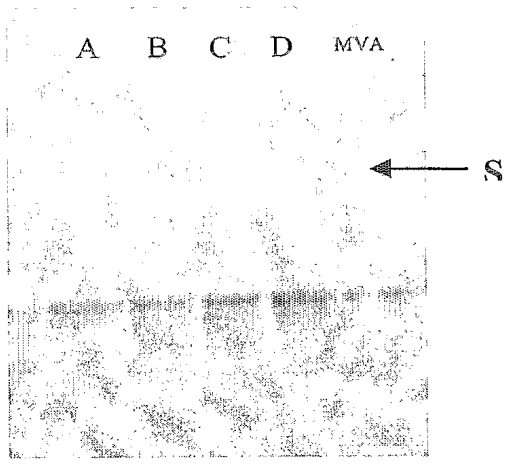

CLONING SARS Spike into pTK53-gpt under VV late P11 promoter
pTK53gpt-Spike
GROWTH AND PURIFICATION of MVA (20)
COTRANSFECTION into Embryo Fibroblast Cells (CEF)
TRANSIENT COTRANSFECTION into CEF
gpt+ SELECTION (2 Passages)
Western/ELISA with SARS Spike Ab
None-selective growth (3 passages)
TK- SELECTION
(24)
Spike gene PCR test
GROWTH AND PURIFICATION of rMVA-Spike (20)
Virus neutralizing antibodies
ANIMAL IMMUNOGENECEITY STUDIES

Figure 62



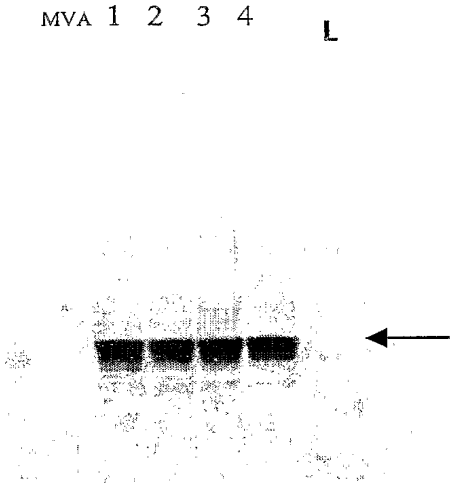
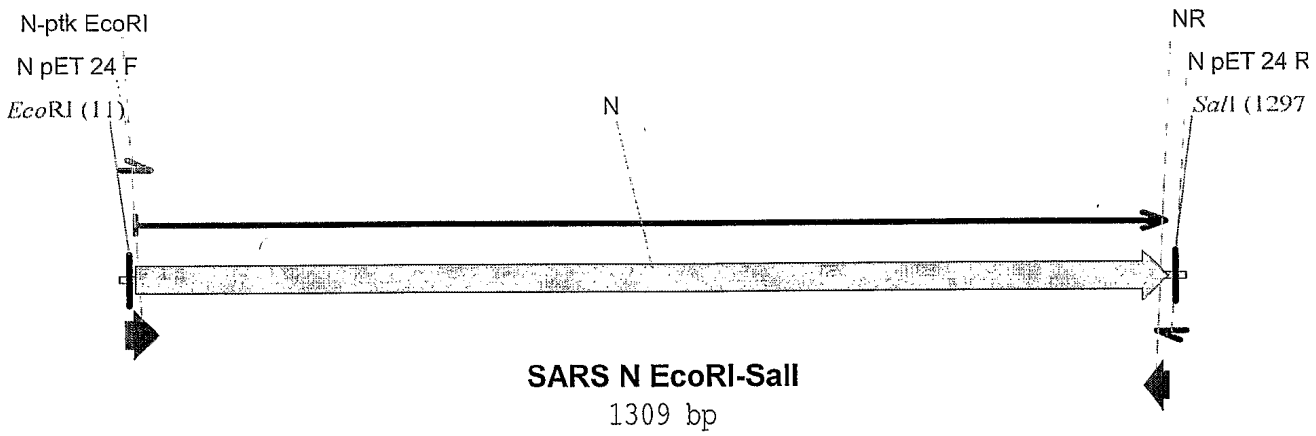


Figure 63

pTK-53 - N

pTK53-N Map



(SEQ ID NO:34)
pTK53-N sequence, from linear map
Start: 19 – End: 1287

atgtctgataatggaccccaatcaaaccaacgtagtgccccccgcattacatttggtggacccacagattcaactgacaataaccagaatgg
aggacgcaatggggcaaggccaaaacagcgccgaccccaaggtttacccaataatactgcgtcttggttcacagctctcactcagcatggc
aaggaggaacttagattccctcgaggccagggcgttccaatcaacaccaatagtggtccagatgaccaaattggctactaccgaagagcta
cccgacgagttcgtggtggtgacggcaaaatgaaagagctcagccccagatggtacttctattacctaggaactggcccagaagcttcactt
ccctacggcgctaacaaagaaggcatcgtatgggttgcaactgagggagccttgaatacacccaaagaccacattggcacccgcaatcct
aataacaatgctgccaccgtgctacaacttcctcaaggaacaacattgccaaaaggcttctacgcagagggaagcagaggcggcagtcaa
gcctcttctcgctcctcatcacgtagtcgcggtaattcaagaaattcaactcctggcagcagtaggggaaattctcctgctcgaatggctagcg
gaggtggtgaaactgccctcgcgctattgctgctagacagattgaaccagcttgagagcaaagtttctggtaaaggccaacaacaacaagg
ccaaactgtcactaagaaatctgctgctgaggcatctaaaaagcctcgccaaaaacgtactgccacaaaacagtacaacgtcactcaagcat
ttgggagacgtggtccagaacaaacccaaggaaatttcggggaccaagacctaatcagacaaggaactgattacaaacattggccgcaaa
ttgcacaatttgctccaagtgcctctgcattctttggaatgtcacgcattggcatggaagtcacaccttcgggaacatggctgacttatcatgga
gccattaaattggatgacaaagatccacaattcaaagacaacgtcatactgctgaacaagcacattgacgcatacaaaacattcccaccaac
agagcctaaaaaggacaaaaagaaaaagactgatgaagctcagcctttgccgcagagacaaaagaagcagcccactgtgactcttcttcct
gcggctgacatggatgatttctccagacaacttcaaaattccatgagtggagcttctgctgattcaactcaggcataa

(SEQ ID NO:35)
pTK 53-N deduced Amino Acid sequence
423 AA

msdngpqsnqrsapritfggptdstdnnqnggrngarpkqrrpqglpnntaswftaltqhgkeelrfprgqgvpintnsgpddqigyy
rratrrvrggdgkmkelsprwyfyylgtgpeaslpygankegivwvategalntpkdhigtrnpnnnaatvlqlpqgttlpkgfyaegs
rggsqassrsssrsrgnsrnstpgssrgnsparmasgggetalallldrlnqleskvsgkgqqqqgqtvtkksaaeaskkprqkrtatkq

ynvtqafgrrgpeqtqgnfgdqdlirqgtdykhwpqiaqfapsasaffgmsrigmevtpsgtwltyhgaiklddkdpqfkdnvillnk
hidayktfpptepkkdkkkktdeaqplpqrqkkqptvtllpaadmddfsrqlqnsmsgasadstqa*

DNA sequence for E2 glycoprotein precursor (Spike glycoprotein)
Length 3,768 nt

SEQ ID NO:36

Start codon to STOP codon

ATGTTTATTTTCTTATTATTTCTTACTCTCACTAGTGGTAGTGACCTTGACCGGTGCACCACTTTTGATGATGTTCAAGC
TCCTAATTACACTCAACATACTTCATCTATGAGGGGGGTTTACTATCCTGATGAAATTTTTAGATCAGACACTCTTTATT
TAACTCAGGATTTATTTCTTCCATTTTATTCTAATGTTACAGGGTTTCATACTATTAATCATACGTTTGGCAACCCTGTC
ATACCTTTTAAGGATGGTATTTATTTTGCTGCCACAGAGAAATCAAATGTTGTCCGTGGTTGGGTTTTTGGTTCTACCAT
GAACAACAAGTCACAGTCGGTGATTATTATTAACAATTCTACTAATGTTGTTATACGAGCATGTAACTTTGAATTGTGTG
ACAACCCTTTCTTTGCTGTTTCTAAACCCATGGGTACACAGACACATACTATGATATTCGATAATGCATTTAATTGCACT
TTCGAGTACATATCTGATGCCTTTTCGCTTGATGTTTCAGAAAAGTCAGGTAATTTTAAACACTTACGAGAGTTTGTGTT
TAAAAATAAAGATGGGTTTCTCTATGTTTATAAGGGCTATCAACCTATAGATGTAGTTCGTGATCTACCTTCTGGTTTTA
ACACTTTGAAACCTATTTTTAAGTTGCCTCTTGGTATTAACATTACAAATTTTAGAGCCATTCTTACAGCCTTTTCACCT
GCTCAAGACATTTGGGGCACGTCAGCTGCAGCCTATTTTGTTGGCTATTTAAAGCCAACTACATTTATGCTCAAGTATGA
TGAAAATGGTACAATCACAGATGCTGTTGATTGTTCTCAAAATCCACTTGCTGAACTCAAATGCTCTGTTAAGAGCTTTG
AGATTGACAAAGGAATTTACCAGACCTCTAATTTCAGGGTTGTTCCCTCAGGAGATGTTGTGAGATTCCCTAATATTACA
AACTTGTGTCCTTTTGGAGAGGTTTTTAATGCTACTAAATTCCCTTCTGTCTATGCATGGGAGAGAAAAAAAATTTCTAA
TTGTGTTGCTGATTACTCTGTGCTCTACAACTCAACATTTTTTTCAACCTTTAAGTGCTATGGCGTTTCTGCCACTAAGT
TGAATGATCTTTGCTTCTCCAATGTCTATGCAGATTCTTTTGTAGTCAAGGGAGATGATGTAAGACAAATAGCGCCAGGA
CAAACTGGTGTTATTGCTGATTATAATTATAAATTGCCAGATGATTTCATGGGTTGTGTCCTTGCTTGGAATACTAGGAA
CATTGATGCTACTTCAACTGGTAATTATAATTATAAATATAGGTATCTTAGACATGGCAAGCTTAGGCCCTTTGAGAGAG
ACATATCTAATGTGCCTTTTCTCCCCTGATGGCAAACCTTGCACCCCACCTGCTCTTAATTGTTATTGGCCATTAAATGAT
TATGGTTTTTACACCACTACTGGCATTGGCTACCAACCTTACAGAGTTGTAGTACTTTCTTTTGAACTTTTAAATGCACC
GGCCACGGTTTGTGGACCAAAATTATCCACTGACCTTATTAAGAACCAGTGTGTCAATTTTAATTTTAATGGACTCACTG
GTACTGGTGTGTTAACTCCTTCTTCAAAGAGATTTCAACCATTTCAACAATTTGGCCGTGATGTTTCTGATTTCACTGAT
TCCGTTCGAGATCCTAAAACATCTGAAATATTAGACATTTCACCTTGCTCTTTTGGGGGTGTAAGTGTAATTACACCTGG
AACAAATGCTTCATCTGAAGTTGCTGTTCTATATCAAGATGTTAACTGCACTGATGTTTCTACAGCAATTCATGCAGATC
AACTCACACCAGCTTGGCGCATATATTCTACTGGAAACAATGTATTCCAGACTCAAGCAGGCTGTCTTATAGGAGCTGAG
CATGTCGACACTTCTTATGAGTGCGACATTCCTATTGGAGCTGGCATTTGTGCTAGTTACCATACAGTTTCTTTATTACG
TAGTACTAGCCAAAAATCTATTGTGGCTTATACTATGTCTTTAGGTGCTGATAGTTCAATTGCTTACTCTAATAACACCA
TTGCTATACCTACTAACTTTTCAATTAGCATTACTACAGAAGTAATGCCTGTTTCTATGGCTAAAACCTCCGTAGATTGT
AATATGTACATCTGCGGAGATTCTACTGAATGTGCTAATTTGCTTCTCCAATATGGTAGCTTTTGCACACAACTAAATCG
TGCACTCTCAGGTATTGCTGCTGAACAGGATCGCAACACACGTGAAGTGTTCGCTCAAGTCAAACAAATGTACAAAACCC
CAACTTTGAAATATTTTGGTGGTTTTAATTTTTCACAAATATTACCTGACCCTCTAAAGCCAACTAAGAGGTCTTTTATT
GAGGACTTGCTCTTTAATAAGGTGACACTCGCTGATGCTGGCTTCATGAAGCAATATGGCGAATGCCTAGGTGATATTAA
TGCTAGAGATCTCATTTGTGCGCAGAAGTTCAATGGACTTACAGTGTTGCCACCTCTGCTCACTGATGATATGATTGCTG
CCTACACTGCTGCTCTAGTTAGTGGTACTGCCACTGCTGGATGGACATTTGGTGCTGGCGCTGCTCTTCAAATACCTTTT
GCTATGCAAATGGCATATAGGTTCAATGGCATTGGAGTTACCCAAAATGTTCTCTATGAGAACCAAAAACAAATCGCCAA
CCAATTTAACAAGGCGATTAGTCAAATTCAAGAATCACTTACAACAACATCAACTGCATTGGGCAAGCTGCAAGACGTTG
TTAACCAGAATGCTCAAGCATTAAACACACTTGTTAAACAACTTAGCTCTAATTTTGGTGCAATTTCAAGTGTGCTAAAT
GATATCCTTTCGCGACTTGATAAAGTCGAGGCGGAGGTACAAATTGACAGGTTAATTACAGGCAGACTTCAAAGCCTTCA
AACCTATGTAACACAACAACTAATCAGGGCTGCTGAAATCAGGGCTTCTGCTAATCTTGCTGCTACTAAAATGTCTGAGT
GTGTTCTTGGACAATCAAAAAGAGTTGACTTTTGTGGAAAGGGCTACCACCTTATGTCCTTCCCACAAGCAGCCCCGCAT
GGTGTTGTCTTCCTACATGTCACGTATGTGCCATCCCAGGAGAGGAACTTCACCACAGCGCCAGCAATTTGTCATGAAGG
CAAAGCATACTTCCCTCGTGAAGGTGTTTTTGTGTTTAATGGCACTTCTTGGTTTATTACACAGAGGAACTTCTTTTCTC
CACAAATAATTACTACAGACAATACATTTGTCTCAGGAAATTGTGATGTCGTTATTGGCATCATTAACAACACAGTTTAT
GATCCTCTGCAACCTGAGCTCGACTCATTCAAAGAAGAGCTGGACAAGTACTTCAAAAATCATACATCACCAGATGTTGA
TCTTGGCGACATTTCAGGCATTAACGCTTCTGTCGTCAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTCGCTAAAA
ATTTAAATGAATCACTCATTGACCTTCAAGAATTGGGAAAATATGAGCAATATATTAAATGGCCTTGGTATGTTTGGCTC
GGCTTCATTGCTGGACTAATTGCCATCGTCATGGTTACAATCTTGCTTTGTTGCATGACTAGTTGTTGCAGTTGCCTCAA
GGGTGCATGCTCTTGTGGTTCTTGCTGCAAGTTTGATGAGGATGACTCTGAGCCAGTTCTCAAGGGTGTCAAATTACATT
ACACATAA

Figure 64 (page 1 of 2)

Protein sequence for E2 glycoprotein precursor (Spike glycoprotein)
Length 1256 aa

Molecular Weight 139,124.54 Daltons

SEQ ID NO:37

MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPV
IPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCT
FEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSP
AQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNIT
NLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPG
QTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLND
YGFYTTTGIGYQPYRVVVLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTD
SVRDPKTSEILDISPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAE
HVDTSYECDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDC
NMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFI
EDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPF
AMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLN
DILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPH
GVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVY
DPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL
GFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT.

Figure 64 (page 2 of 2)

Contig0001.TXT

The entire nucleotide sequence of SARS-CoV (Urbani strain).
The genome is 29, 727 nucleotides in length from 5' leader to 3'end.

Contig[0001] Length: 29727 Mon, Apr 14, 2003 10:55 AM Check: 4614

```
   1  TTATTAGGTT TTTACCTACC CAGGAAAAGC CAACCAACCT CGATCTCTTG
  51  TAGATCTGTT CTCTAAACGA ACTTTAAAAT CTGTGTAGCT GTCGCTCGGC
 101  TGCATGCCTA GTGCACCTAC GCAGTATAAA CAATAATAAA TTTTACTGTC
 151  GTTGACAAGA AACGAGTAAC TCGTCCCTCT TCTGCAGACT GCTTACGGTT
 201  TCGTCCGTGT TGCAGTCGAT CATCAGCATA CCTAGGTTTC GTCCGGGTGT
 251  GACCGAAAGG TAAGATGGAG AGCCTTGTTC TTGGTGTCAA CGAGAAAACA
 301  CACGTCCAAC TCAGTTTGCC TGTCCTTCAG GTTAGAGACG TGCTAGTGCG
 351  TGGCTTCGGG GACTCTGTGG AAGAGGCCCT ATCGGAGGCA CGTGAACACC
 401  TCAAAAATGG CACTTGTGGT CTAGTAGAGC TGGAAAAAGG CGTACTGCCC
 451  CAGCTTGAAC AGCCCTATGT GTTCATTAAA CGTTCTGATG CCTTAAGCAC
 501  CAATCACGGC CACAAGGTCG TTGAGCTGGT TGCAGAAATG GACGGCATTC
 551  AGTACGGTCG TAGCGGTATA ACACTGGGAG TACTCGTGCC ACATGTGGGC
 601  GAAACCCCAA TTGCATACCG CAATGTTCTT CTTCGTAAGA ACGGTAATAA
 651  GGGAGCCGGT GGTCATAGCT ATGGCATCGA TCTAAAGTCT TATGACTTAG
 701  GTGACGAGCT TGGCACTGAT CCCATTGAAG ATTATGAACA AAACTGGAAC
 751  ACTAAGCATG GCAGTGGTGC ACTCCGTGAA CTCACTCGTG AGCTCAATGG
 801  AGGTGCAGTC ACTCGCTATG TCGACAACAA TTTCTGTGGC CCAGATGGGT
 851  ACCCTCTTGA TTGCATCAAA GATTTTCTCG CACGCGCGGG CAAGTCAATG
 901  TGCACTCTTT CCGAACAACT TGATTACATC GAGTCGAAGA GAGGTGTCTA
 951  CTGCTGCCGT GACCATGAGC ATGAAATTGC CTGGTTCACT GAGCGCTCTG
1001  ATAAGAGCTA CGAGCACCAG ACACCCTTCG AAATTAAGAG TGCCAAGAAA
1051  TTTGACACTT TCAAAGGGGA ATGCCCAAAG TTTGTGTTTC CTCTTAACTC
1101  AAAAGTCAAA GTCATTCAAC CACGTGTTGA AAAGAAAAAG ACTGAGGGTT
1151  TCATGGGGCG TATACGCTCT GTGTACCCTG TTGCATCTCC ACAGGAGTGT
1201  AACAATATGC ACTTGTCTAC CTTGATGAAA TGTAATCATT GCGATGAAGT
1251  TTCATGGCAG ACGTGCGACT TTCTGAAAGC CACTTGTGAA CATTGTGGCA
1301  CTGAAAATTT AGTTATTGAA GGACCTACTA CATGTGGGTA CCTACCTACT
1351  AATGCTGTAG TGAAAATGCC ATGTCCTGCC TGTCAAGACC CAGAGATTGG
1401  ACCTGAGCAT AGTGTTGCAG ATTATCACAA CCACTCAAAC ATTGAAACTC
1451  GACTCCGCAA GGGAGGTAGG ACTAGATGTT TTGGAGGCTG TGTGTTTGCC
1501  TATGTTGGCT GCTATAATAA GCGTGCCTAC TGGGTTCCTC GTGCTAGTGC
1551  TGATATTGGC TCAGGCCATA CTGGCATTAC TGGTGACAAT GTGGAGACCT
1601  TGAATGAGGA TCTCCTTGAG ATACTGAGTC GTGAACGTGT TAACATTAAC
1651  ATTGTTGGCG ATTTTCATTT GAATGAAGAG GTTGCCATCA TTTTGGCATC
1701  TTTCTCTGCT TCTACAAGTG CCTTTATTGA CACTATAAAG AGTCTTGATT
1751  ACAAGTCTTT CAAAACCATT GTTGAGTCCT GCGGTAACTA TAAAGTTACC
1801  AAGGGAAAGC CCGTAAAAGG TGCTTGGAAC ATTGGACAAC AGAGATCAGT
1851  TTTAACACCA CTGTGTGGTT TTCCCTCACA GGCTGCTGGT GTTATCAGAT
1901  CAATTTTTGC GCGCACACTT GATGCAGCAA ACCACTCAAT TCCTGATTTG
1951  CAAAGAGCAG CTGTCACCAT ACTTGATGGT ATTTCTGAAC AGTCATTACG
2001  TCTTGTCGAC GCCATGGTTT ATACTTCAGA CCTGCTCACC AACAGTGTCA
2051  TTATTATGGC ATATGTAACT GGTGGTCTTG TACAACAGAC TTCTCAGTGG
2101  TTGTCTAATC TTTTGGGCAC TACTGTTGAA AAACTCAGGC CTATCTTTGA
2151  ATGGATTGAG GCGAAACTTA GTGCAGGAGT TGAATTTCTC AAGGATGCTT
2201  GGGAGATTCT CAAATTTCTC ATTACAGGTG TTTTTGACAT CGTCAAGGGT
2251  CAAATACAGG TTGCTTCAGA TAACATCAAG GATTGTGTAA AATGCTTCAT
2301  TGATGTTGTT AACAAGGCAC TCGAAATGTG CATTGATCAA GTCACTATCG
2351  CTGGCGCAAA GTTGCGATCA CTCAACTTAG GTGAAGTCTT CATCGCTCAA
2401  AGCAAGGGAC TTTACCGTCA GTGTATACGT GGCAAGGAGC AGCTGCAACT
2451  ACTCATGCCT CTTAAGGCAC CAAAAGAAGT AACCTTTCTT GAAGGTGATT
2501  CACATGACAC AGTACTTACC TCTGAGGAGG TTGTTCTCAA GAACGGTGAA
2551  CTCGAAGCAC TCGAGACGCC CGTTGATAGC TTCACAAATG GAGCTATCGT
2601  TGGCACACCA GTCTGTGTAA ATGGCCTCAT GCTCTTAGAG ATTAAGGACA
2651  AAGAACAATA CTGCGCATTG TCTCCTGGTT TACTGGCTAC AAACAATGTC
2701  TTTCGCTTAA AAGGGGGTGC ACCAATTAAA GGTGTAACCT TTGGAGAAGA
2751  TACTGTTTGG GAAGTTCAAG GTTACAAGAA TGTGAGAATC ACATTTGAGC
2801  TTGATGAACG TGTTGACAAA GTGCTTAATG AAAAGTGCTC TGTCTACACT
```

Figure 65 (page 1 of 10)

Contig0001.TXT

```
2851  GTTGAATCCG GTACCGAAGT TACTGAGTTT GCATGTGTTG TAGCAGAGGC
2901  TGTTGTGAAG ACTTTACAAC CAGTTTCTGA TCTCCTTACC AACATGGGTA
2951  TTGATCTTGA TGAGTGGAGT GTAGCTACAT TCTACTTATT TGATGATGCT
3001  GGTGAAGAAA ACTTTTCATC ACGTATGTAT TGTTCCTTTT ACCCTCCAGA
3051  TGAGGAAGAA GAGGACGATG CAGAGTGTGA GGAAGAAGAA ATTGATGAAA
3101  CCTGTGAACA TGAGTACGGT ACAGAGGATG ATTATCAAGG TCTCCCTCTG
3151  GAATTTGGTG CCTCAGCTGA AACAGTTCGA GTTGAGGAAG AAGAAGAGGA
3201  AGACTGGCTG GATGATACTA CTGAGCAATC AGAGATTGAG CCAGAACCAG
3251  AACCTACACC TGAAGAACCA GTTAATCAGT TTACTGGTTA TTTAAAACTT
3301  ACTGACAATG TTGCCATTAA ATGTGTTGAC ATCGTTAAGG AGGCACAAAG
3351  TGCTAATCCT ATGGTGATTG TAAATGCTGC TAACATACAC CTGAAACATG
3401  GTGGTGGTGT AGCAGGTGCA CTCAACAAGG CAACCAATGG TGCCATGCAA
3451  AAGGAGAGTG ATGATTACAT TAAGCTAAAT GGCCCTCTTA CAGTAGGAGG
3501  GTCTTGTTTG CTTTCTGGAC ATAATCTTGC TAAGAAGTGT CTGCATGTTG
3551  TTGGACCTAA CCTAAATGCA GGTGAGGACA TCCAGCTTCT TAAGGCAGCA
3601  TATGAAAATT TCAATTCACA GGACATCTTA CTTGCACCAT TGTTGTCAGC
3651  AGGCATATTT GGTGCTAAAC CACTTCAGTC TTTACAAGTG TGCGTGCAGA
3701  CGGTTCGTAC ACAGGTTTAT ATTGCAGTCA ATGACAAAGC TCTTTATGAG
3751  CAGGTTGTCA TGGATTATCT TGATAACCTG AAGCCTAGAG TGGAAGCACC
3801  TAAACAAGAG GAGCCACCAA ACACAGAAGA TTCCAAAACT GAGGAGAAAT
3851  CTGTCGTACA GAAGCCTGTC GATGTGAAGC CAAAAATTAA GGCCTGCATT
3901  GATGAGGTTA CCACAACACT GGAAGAAACT AAGTTTCTTA CCAATAAGTT
3951  ACTCTTGTTT GCTGATATCA ATGGTAAGCT TTACCATGAT TCTCAGAACA
4001  TGCTTAGAGG TGAAGATATG TCTTTCCTTG AGAAGGATGC ACCTTACATG
4051  GTAGGTGATG TTATCACTAG TGGTGATATC ACTTGTGTTG TAATACCCTC
4101  CAAAAAGGCT GGTGGCACTA CTGAGATGCT CTCAAGAGCT TTGAAGAAAG
4151  TGCCAGTTGA TGAGTATATA ACCACGTACC CTGGACAAGG ATGTGCTGGT
4201  TATACACTTG AGGAAGCTAA GACTGCTCTT AAGAAATGCA AATCTGCATT
4251  TTATGTACTA CCTTCAGAAG CACCTAATGC TAAGGAAGAG ATTCTAGGAA
4301  CTGTATCCTG GAATTTGAGA GAAATGCTTG CTCATGCTGA AGAGACAAGA
4351  AAATTAATGC CTATATGCAT GGATGTTAGA GCCATAATGG CAACCATCCA
4401  ACGTAAGTAT AAAGGAATTA AAATTCAAGA GGGCATCGTT GACTATGGTG
4451  TCCGATTCTT CTTTTATACT AGTAAAGAGC CTGTAGCTTC TATTATTACG
4501  AAGCTGAACT CTCTAAATGA GCCGCTTGTC ACAATGCCAA TTGGTTATGT
4551  GACACATGGT TTTAATCTTG AAGAGGCTGC GCGCTGTATG CGTTCTCTTA
4601  AAGCTCCTGC CGTAGTGTCA GTATCATCAC CAGATGCTGT TACTACATAT
4651  AATGGATACC TCACTTCGTC ATCAAAGACA TCTGAGGAGC ACTTTGTAGA
4701  AACAGTTTCT TTGGCTGGCT CTTACAGAGA TTGGTCCTAT TCAGGACAGC
4751  GTACAGAGTT AGGTGTTGAA TTTCTTAAGC GTGGTGACAA AATTGTGTAC
4801  CACACTCTGG AGAGCCCCGT CGAGTTTCAT CTTGACGGTG AGGTTCTTTC
4851  ACTTGACAAA CTAAAGAGTC TCTTATCCCT GCGGGAGGTT AAGACTATAA
4901  AAGTGTTCAC AACTGTGGAC AACACTAATC TCCACACACA GCTTGTGGAT
4951  ATGTCTATGA CATATGGACA GCAGTTTGGT CCAACATACT TGGATGGTGC
5001  TGATGTTACA AAAATTAAAC CTCATGTAAA TCATGAGGGT AAGACTTTCT
5051  TTGTACTACC TAGTGATGAC ACACTACGTA GTGAAGCTTT CGAGTACTAC
5101  CATACTCTTG ATGAGAGTTT TCTTGGTAGG TACATGTCTG CTTTAAACCA
5151  CACAAAGAAA TGGAAATTTC CTCAAGTTGG TGGTTTAACT TCAATTAAAT
5201  GGGCTGATAA CAATTGTTAT TTGTCTAGTG TTTTATTAGC ACTTCAACAG
5251  CTTGAAGTCA AATTCAATGC ACCAGCACTT CAAGAGGCTT ATTATAGAGC
5301  CCGTGCTGGT GATGCTGCTA ACTTTTGTGC ACTCATACTC GCTTACAGTA
5351  ATAAAACTGT TGGCGAGCTT GGTGATGTCA GAGAAACTAT GACCCATCTT
5401  CTACAGCATG CTAATTTGGA ATCTGCAAAG CGAGTTCTTA ATGTGGTGTG
5451  TAAACATTGT GGTCAGAAAA CTACTACCTT AACGGGTGTA GAAGCTGTGA
5501  TGTATATGGG TACTCTATCT TATGATAATC TTAAGACAGG TGTTTCCATT
5551  CCATGTGTGT GTGGTCGTGA TGCTACACAA TATCTAGTAC AACAAGAGTC
5601  TTCTTTTGTT ATGATGTCTG CACCACCTGC TGAGTATAAA TTACAGCAAG
5651  GTACATTCTT ATGTGCGAAT GAGTACACTG GTAACTATCA GTGTGGTCAT
5701  TACACTCATA TAACTGCTAA GGAGACCCTC TATCGTATTG ACGGAGCTCA
5751  CCTTACAAAG ATGTCAGAGT ACAAAGGACC AGTGACTGAT GTTTTCTACA
5801  AGGAAACATC TTACACTACA ACCATCAAGC CTGTGTCGTA TAAACTCGAT
5851  GGAGTTACTT ACACAGAGAT TGAACCAAAA TTGGATGGGT ATTATAAAAA
5901  GGATAATGCT TACTATACAG AGCAGCCTAT AGACCTTGTA CCAACTCAAC
5951  CATTACCAAA TGCGAGTTTT GATAATTTCA AACTCACATG TTCTAACACA
```

Figure 65 (page 2 of 10)

Contig0001.TXT

6001	AAATTTGCTG	ATGATTTAAA	TCAAATGACA	GGCTTCACAA	AGCCAGCTTC
6051	ACGAGAGCTA	TCTGTCACAT	TCTTCCCAGA	CTTGAATGGC	GATGTAGTGG
6101	CTATTGACTA	TAGACACTAT	TCAGCGAGTT	TCAAGAAAGG	TGCTAAATTA
6151	CTGCATAAGC	CAATTGTTTG	GCACATTAAC	CAGGCTACAA	CCAAGACAAC
6201	GTTCAAACCA	AACACTTGGT	GTTTACGTTG	TCTTTGGAGT	ACAAAGCCAG
6251	TAGATACTTC	AAATTCATTT	GAAGTTCTGG	CAGTAGAAGA	CACACAAGGA
6301	ATGGACAATC	TTGCTTGTGA	AAGTCAACAA	CCCACCTCTG	AAGAAGTAGT
6351	GGAAAATCCT	ACCATACAGA	AGGAAGTCAT	AGAGTGTGAC	GTGAAAACTA
6401	CCGAAGTTGT	AGGCAATGTC	ATACTTAAAC	CATCAGATGA	AGGTGTTAAA
6451	GTAACACAAG	AGTTAGGTCA	TGAGGATCTT	ATGGCTGCTT	ATGTGGAAAA
6501	CACAAGCATT	ACCATTAAGA	AACCTAATGA	GCTTTCACTA	GCCTTAGGTT
6551	TAAAAACAAT	TGCCACTCAT	GGTATTGCTG	CAATTAATAG	TGTTCCTTGG
6601	AGTAAAATTT	TGGCTTATGT	CAAACCATTC	TTAGGACAAG	CAGCAATTAC
6651	AACATCAAAT	TGCGCTAAGA	GATTAGCACA	ACGTGTGTTT	AACAATTATA
6701	TGCCTTATGT	GTTTACATTA	TTGTTCCAAT	TGTGTACTTT	TACTAAAAGT
6751	ACCAATTCTA	GAATTAGAGC	TTCACTACCT	ACAACTATTG	CTAAAAATAG
6801	TGTTAAGAGT	GTTGCTAAAT	TATGTTTGGA	TGCCGGCATT	AATTATGTGA
6851	AGTCACCCAA	ATTTTCTAAA	TTGTTCACAA	TCGCTATGTG	GCTATTGTTG
6901	TTAAGTATTT	GCTTAGGTTC	TCTAATCTGT	GTAACTGCTG	CTTTTGGTGT
6951	ACTCTTATCT	AATTTTGGTG	CTCCTTCTTA	TTGTAATGGC	GTTAGAGAAT
7001	TGTATCTTAA	TTCGTCTAAC	GTTACTACTA	TGGATTTCTG	TGAAGGTTCT
7051	TTTCCTTGCA	GCATTTGTTT	AAGTGGATTA	GACTCCCTTG	ATTCTTATCC
7101	AGCTCTTGAA	ACCATTCAGG	TGACGATTTC	ATCGTACAAG	CTAGACTTGA
7151	CAATTTTAGG	TCTGGCCGCT	GAGTGGGTTT	TGGCATATAT	GTTGTTCACA
7201	AAATTCTTTT	ATTTATTAGG	TCTTTCAGCT	ATAATGCAGG	TGTTCTTTGG
7251	CTATTTTGCT	AGTCATTTCA	TCAGCAATTC	TTGGCTCATG	TGGTTTATCA
7301	TTAGTATTGT	ACAAATGGCA	CCCGTTTCTG	CAATGGTTAG	GATGTACATC
7351	TTCTTTGCTT	CTTTCTACTA	CATATGGAAG	AGCTATGTTC	ATATCATGGA
7401	TGGTTGCACC	TCTTCGACTT	GCATGATGTG	CTATAAGCGC	AATCGTGCCA
7451	CACGCGTTGA	GTGTACAACT	ATTGTTAATG	GCATGAAGAG	ATCTTTCTAT
7501	GTCTATGCAA	ATGGAGGCCG	TGGCTTCTGC	AAGACTCACA	ATTGGAATTG
7551	TCTCAATTGT	GACACATTTT	GCACTGGTAG	TACATTCATT	AGTGATGAAG
7601	TTGCTCGTGA	TTTGTCACTC	CAGTTTAAAA	GACCAATCAA	CCCTACTGAC
7651	CAGTCATCGT	ATATTGTTGA	TAGTGTTGCT	GTGAAAAATG	GCGCGCTTCA
7701	CCTCTACTTT	GACAAGGCTG	GTCAAAAGAC	CTATGAGAGA	CATCCGCTCT
7751	CCCATTTTGT	CAATTTAGAC	AATTTGAGAG	CTAACAACAC	TAAAGGTTCA
7801	CTGCCTATTA	ATGTCATAGT	TTTTGATGGC	AAGTCCAAAT	GCGACGAGTC
7851	TGCTTCTAAG	TCTGCTTCTG	TGTACTACAG	TCAGCTGATG	TGCCAACCTA
7901	TTCTGTTGCT	TGACCAAGTT	CTTGTATCAG	ACGTTGGAGA	TAGTACTGAA
7951	GTTTCCGTTA	AGATGTTTGA	TGCTTATGTC	GACACCTTTT	CAGCAACTTT
8001	TAGTGTTCCT	ATGGAAAAAC	TTAAGGCACT	TGTTGCTACA	GCTCACAGCG
8051	AGTTAGCAAA	GGGTGTAGCT	TTAGATGGTG	TCCTTTCTAC	ATTCGTGTCA
8101	GCTGCCCGAC	AAGGTGTTGT	TGATACCGAT	GTTGACACAA	AGGATGTTAT
8151	TGAATGTCTC	AAACTTTCAC	ATCACTCTGA	CTTAGAAGTG	ACAGGTGACA
8201	GTTGTAACAA	TTTCATGCTC	ACCTATAATA	AGGTTGAAAA	CATGACGCCC
8251	AGAGATCTTG	GCGCATGTAT	TGACTGTAAT	GCAAGGCATA	TCAATGCCCA
8301	AGTAGCAAAA	AGTCACAATG	TTTCACTCAT	CTGGAATGTA	AAAGACTACA
8351	TGTCTTTATC	TGAACAGCTG	CGTAAACAAA	TTCGTAGTGC	TGCCAAGAAG
8401	AACAACATAC	CTTTTAGACT	AACTTGTGCT	ACAACTAGAC	AGGTTGTCAA
8451	TGTCATAACT	ACTAAAATCT	CACTCAAGGG	TGGTAAGATT	GTTAGTACTT
8501	GTTTTAAACT	TATGCTTAAG	GCCACATTAT	TGTGCGTTCT	TGCTGCATTG
8551	GTTTGTTATA	TCGTTATGCC	AGTACATACA	TTGTCAATCC	ATGATGGTTA
8601	CACAAATGAA	ATCATTGGTT	ACAAAGCCAT	TCAGGATGGT	GTCACTCGTG
8651	ACATCATTTC	TACTGATGAT	TGTTTTGCAA	ATAAACATGC	TGGTTTTGAC
8701	GCATGGTTTA	GCCAGCGTGG	TGGTTCATAC	AAAAATGACA	AAAGCTGCCC
8751	TGTAGTAGCT	GCTATCATTA	CAAGAGAGAT	TGGTTTCATA	GTGCCTGGCT
8801	TACCGGGTAC	TGTGCTGAGA	GCAATCAATG	GTGACTTCTT	GCATTTTCTA
8851	CCTCGTGTTT	TTAGTGCTGT	TGGCAACATT	TGCTACACAC	CTTCCAAACT
8901	CATTGAGTAT	AGTGATTTTG	CTACCTCTGC	TTGCGTTCTT	GCTGCTGAGT
8951	GTACAATTTT	TAAGGATGCT	ATGGGCAAAC	CTGTGCCATA	TTGTTATGAC
9001	ACTAATTTGC	TAGAGGGTTC	TATTTCTTAT	AGTGAGCTTC	GTCCAGACAC
9051	TCGTTATGTG	CTTATGGATG	GTTCCATCAT	ACAGTTTCCT	AACACTTACC
9101	TGGAGGGTTC	TGTTAGAGTA	GTAACAACTT	TTGATGCTGA	GTACTGTAGA

Figure 65 (page 3 of 10)

Contig0001.TXT

```
 9151  CATGGTACAT GCGAAAGGTC AGAAGTAGGT ATTTGCCTAT CTACCAGTGG
 9201  TAGATGGGTT CTTAATAATG AGCATTACAG AGCTCTATCA GGAGTTTTCT
 9251  GTGGTGTTGA TGCGATGAAT CTCATAGCTA ACATCTTTAC TCCTCTTGTG
 9301  CAACCTGTGG GTGCTTTAGA TGTGTCTGCT TCAGTAGTGG CTGGTGGTAT
 9351  TATTGCCATA TTGGTGACTT GTGCTGCCTA CTACTTTATG AAATTCAGAC
 9401  GTGTTTTTGG TGAGTACAAC CATGTTGTTG CTGCTAATGC ACTTTTGTTT
 9451  TTGATGTCTT TCACTATACT CTGTCTGGTA CCAGCTTACA GCTTTCTGCC
 9501  GGGAGTCTAC TCAGTCTTTT ACTTGTACTT GACATTCTAT TTCACCAATG
 9551  ATGTTTCATT CTTGGCTCAC CTTCAATGGT TTGCCATGTT TTCTCCTATT
 9601  GTGCCTTTTT GGATAACAGC AATCTATGTA TTCTGTATTT CTCTGAAGCA
 9651  CTGCCATTGG TTCTTTAACA ACTATCTTAG GAAAAGAGTC ATGTTTAATG
 9701  GAGTTACATT TAGTACCTTC GAGGAGGCTG CTTTGTGTAC CTTTTTGCTC
 9751  AACAAGGAAA TGTACCTAAA ATTGCGTAGC GAGACACTGT TGCCACTTAC
 9801  ACAGTATAAC AGGTATCTTG CTCTATATAA CAAGTACAAG TATTTCAGTG
 9851  GAGCCTTAGA TACTACCAGC TATCGTGAAG CAGCTTGCTG CCACTTAGCA
 9901  AAGGCTCTAA ATGACTTTAG CAACTCAGGT GCTGATGTTC TCTACCAACC
 9951  ACCACAGACA TCAATCACTT CTGCTGTTCT GCAGAGTGGT TTTAGGAAAA
10001  TGGCATTCCC GTCAGGCAAA GTTGAAGGGT GCATGGTACA AGTAACCTGT
10051  GGAACTACAA CTCTTAATGG ATTGTGGTTG GATGACACAG TATACTGTCC
10101  AAGACATGTC ATTTGCACAG CAGAAGACAT GCTTAATCCT AACTATGAAG
10151  ATCTGCTCAT TCGCAAATCC AACCATAGCT TTCTTGTTCA GGCTGGCAAT
10201  GTTCAACTTC GTGTTATTGG CCATTCTATG CAAAATTGTC TGCTTAGGCT
10251  TAAAGTTGAT ACTTCTAACC CTAAGACACC CAAGTATAAA TTTGTCCGTA
10301  TCCAACCTGG TCAAACATTT TCAGTTCTAG CATGCTACAA TGGTTCACCA
10351  TCTGGTGTTT ATCAGTGTGC CATGAGACCT AATCATACCA TTAAAGGTTC
10401  TTTCCTTAAT GGATCATGTG GTAGTGTTGG TTTTAACATT GATTATGATT
10451  GCGTGTCTTT CTGCTATATG CATCATATGG AGCTTCCAAC AGGAGTACAC
10501  GCTGGTACTG ACTTAGAAGG TAAATTCTAT GGTCCATTTG TTGACAGACA
10551  AACTGCACAG GCTGCAGGTA CAGACACAAC CATAACATTA AATGTTTTGG
10601  CATGGCTGTA TGCTGCTGTT ATCAATGGTG ATAGGTGGTT TCTTAATAGA
10651  TTCACCACTA CTTTGAATGA CTTTAACCTT GTGGCAATGA AGTACAACTA
10701  TGAACCTTTG ACACAAGATC ATGTTGACAT ATTGGGACCT CTTTCTGCTC
10751  AAACAGGAAT TGCCGTCTTA GATATGTGTG CTGCTTTGAA AGAGCTGCTG
10801  CAGAATGGTA TGAATGGTCG TACTATCCTT GGTAGCACTA TTTTAGAAGA
10851  TGAGTTTACA CCATTTGATG TTGTTAGACA ATGCTCTGGT GTTACCTTCC
10901  AAGGTAAGTT CAAGAAAATT GTTAAGGGCA CTCATCATTG GATGCTTTTA
10951  ACTTTCTTGA CATCACTATT GATTCTTGTT CAAAGTACAC AGTGGTCACT
11001  GTTTTTCTTT GTTTACGAGA ATGCTTTCTT GCCATTTACT CTTGGTATTA
11051  TGGCAATTGC TGCATGTGCT ATGCTGCTTG TTAAGCATAA GCACGCATTC
11101  TTGTGCTTGT TTCTGTTACC TTCTCTTGCA ACAGTTGCTT ACTTTAATAT
11151  GGTCTACATG CCTGCTAGCT GGGTGATGCG TATCATGACA TGGCTTGAAT
11201  TGGCTGACAC TAGCTTGTCT GGTTATAGGC TTAAGGATTG TGTTATGTAT
11251  GCTTCAGCTT TAGTTTTGCT TATTCTCATG ACAGCTCGCA CTGTTTATGA
11301  TGATGCTGCT AGACGTGTTT GGACACTGAT GAATGTCATT ACACTTGTTT
11351  ACAAAGTCTA CTATGGTAAT GCTTTAGATC AAGCTATTTC CATGTGGGCC
11401  TTAGTTATTT CTGTAACCTC TAACTATTCT GGTGTCGTTA CGACTATCAT
11451  GTTTTTAGCT AGAGCTATAG TGTTTGTGTG TGTTGAGTAT TACCCATTGT
11501  TATTTATTAC TGGCAACACC TTACAGTGTA TCATGCTTGT TTATTGTTTC
11551  TTAGGCTATT GTTGCTGCTG CTACTTTGGC CTTTTCTGTT TACTCAACCG
11601  TTACTTCAGG CTTACTCTTG GTGTTTATGA CTACTTGGTC TCTACACAAG
11651  AATTTAGGTA TATGAACTCC CAGGGGCTTT TGCCTCCTAA GAGTAGTATT
11701  GATGCTTTCA AGCTTAACAT TAAGTTGTTG GGTATTGGAG GTAAACCATG
11751  TATCAAGGTT GCTACTGTAC AGTCTAAAAT GTCTGACGTA AAGTGCACAT
11801  CTGTGGTACT GCTCTCGGTT CTTCAACAAC TTAGAGTAGA GTCATCTTCT
11851  AAATTGTGGG CACAATGTGT ACAACTCCAC AATGATATTC TTCTTGCAAA
11901  AGACACAACT GAAGCTTTCG AGAAGATGGT TTCTCTTTTG TCTGTTTTGC
11951  TATCCATGCA GGGTGCTGTA GACATTAATA GGTTGTGCGA GGAAATGCTC
12001  GATAACCGTG CTACTCTTCA GGCTATTGCT TCAGAATTTA GTTCTTTACC
12051  ATCATATGCC GCTTATGCCA CTGCCCAGGA GGCCTATGAG CAGGCTGTAG
12101  CTAATGGTGA TTCTGAAGTC GTTCTCAAAA AGTTAAAGAA ATCTTTGAAT
12151  GTGGCTAAAT CTGAGTTTGA CCGTGATGCT GCCATGCAAC GCAAGTTGGA
12201  AAAGATGGCA GATCAGGCTA TGACCCAAAT GTACAAACAG GCAAGATCTG
12251  AGGACAAGAG GGCAAAAGTA ACTAGTGCTA TGCAAACAAT GCTCTTCACT
```

Figure 65 (page 4 of 10)

Contig0001.TXT

```
12301  ATGCTTAGGA AGCTTGATAA TGATGCACTT AACAACATTA TCAACAATGC
12351  GCGTGATGGT TGTGTTCCAC TCAACATCAT ACCATTGACT ACAGCAGCCA
12401  AACTCATGGT TGTTGTCCCT GATTATGGTA CCTACAAGAA CACTTGTGAT
12451  GGTAACACCT TTACATATGC ATCTGCACTC TGGGAAATCC AGCAAGTTGT
12501  TGATGCGGAT AGCAAGATTG TTCAACTTAG TGAAATTAAC ATGGACAATT
12551  CACCAAATTT GGCTTGGCCT CTTATTGTTA CAGCTCTAAG AGCCAACTCA
12601  GCTGTTAAAC TACAGAATAA TGAACTGAGT CCAGTAGCAC TACGACAGAT
12651  GTCCTGTGCG GCTGGTACCA CACAAACAGC TTGTACTGAT GACAATGCAC
12701  TTGCCTACTA TAACAATTCG AAGGGAGGTA GGTTTGTGCT GGCATTACTA
12751  TCAGACCACC AAGATCTCAA ATGGGCTAGA TTCCCTAAGA GTGATGGTAC
12801  AGGTACAATT TACACAGAAC TGGAACCACC TTGTAGGTTT GTTACAGACA
12851  CACCAAAAGG GCCTAAAGTG AAATACTTGT ACTTCATCAA AGGCTTAAAC
12901  AACCTAAATA GAGGTATGGT GCTGGGCAGT TTAGCTGCTA CAGTACGTCT
12951  TCAGGCTGGA AATGCTACAG AAGTACCTGC CAATTCAACT GTGCTTTCCT
13001  TCTGTGCTTT TGCAGTAGAC CCTGCTAAAG CATATAAGGA TTACCTAGCA
13051  AGTGGAGGAC AACCAATCAC CAACTGTGTG AAGATGTTGT GTACACACAC
13101  TGGTACAGGA CAGGCAATTA CTGTAACACC AGAAGCTAAC ATGGACCAAG
13151  AGTCCTTTGG TGGTGCTTCA TGTTGTCTGT ATTGTAGATG CCACATTGAC
13201  CATCCAAATC CTAAAGGATT CTGTGACTTG AAAGGTAAGT ACGTCCAAAT
13251  ACCTACCACT TGTGCTAATG ACCCAGTGGG TTTTACACTT AGAAACACAG
13301  TCTGTACCGT CTGCGGAATG TGGAAAGGTT ATGGCTGTAG TTGTGACCAA
13351  CTCCGCGAAC CCTTGATGCA GTCTGCGGAT GCATCAACGT TTTTAAACGG
13401  GTTTGCGGTG TAAGTGCAGC CCGTCTTACA CCGTGCGGCA CAGGCACTAG
13451  TACTGATGTC GTCTACAGGG CTTTTGATAT TTACAACGAA AAAGTTGCTG
13501  GTTTTGCAAA GTTCCTAAAA ACTAATTGCT GTCGCTTCCA GGAGAAGGAT
13551  GAGGAAGGCA ATTTATTAGA CTCTTACTTT GTAGTTAAGA GGCATACTAT
13601  GTCTAACTAC CAACATGAAG AGACTATTTA TAACTTGGTT AAAGATTGTC
13651  CAGCGGTTGC TGTCCATGAC TTTTTCAAGT TTAGAGTAGA TGGTGACATG
13701  GTACCACATA TATCACGTCA GCGTCTAACT AAATACACAA TGGCTGATTT
13751  AGTCTATGCT CTACGTCATT TTGATGAGGG TAATTGTGAT ACATTAAAAG
13801  AAATACTCGT CACATACAAT TGCTGTGATG ATGATTATTT CAATAAGAAG
13851  GATTGGTATG ACTTCGTAGA GAATCCTGAC ATCTTACGCG TATATGCTAA
13901  CTTAGGTGAG CGTGTACGCC AATCATTATT AAAGACTGTA CAATTCTGCG
13951  ATGCTATGCG TGATGCAGGC ATTGTAGGCG TACTGACATT AGATAATCAG
14001  GATCTTAATG GGAACTGGTA CGATTTCGGT GATTTCGTAC AAGTAGCACC
14051  AGGCTGCGGA GTTCCTATTG TGGATTCATA TTACTCATTG CTGATGCCCA
14101  TCCTCACTTT GACTAGGGCA TTGGCTGCTG AGTCCCATAT GGATGCTGAT
14151  CTCGCAAAAC CACTTATTAA GTGGGATTTG CTGAAATATG ATTTTACGGA
14201  AGAGAGACTT TGTCTCTTCG ACCGTTATTT TAAATATTGG GACCAGACAT
14251  ACCATCCCAA TTGTATTAAC TGTTTGGATG ATAGGTGTAT CCTTCATTGT
14301  GCAAACTTTA ATGTGTTATT TTCTACTGTG TTTCCACCTA CAAGTTTTGG
14351  ACCACTAGTA AGAAAAATAT TTGTAGATGG TGTTCCTTTT GTTGTTTCAA
14401  CTGGATACCA TTTTCGTGAG TTAGGAGTCG TACATAATCA GGATGTAAAC
14451  TTACATAGCT CGCGTCTCAG TTTCAAGGAA CTTTTAGTGT ATGCTGCTGA
14501  TCCAGCTATG CATGCAGCTT CTGGCAATTT ATTGCTAGAT AAACGCACTA
14551  CATGCTTTTC AGTAGCTGCA CTAACAAACA ATGTTGCTTT TCAAACTGTC
14601  AAACCCGGTA ATTTTAATAA AGACTTTTAT GACTTTGCTG TGTCTAAAGG
14651  TTTCTTTAAG GAAGGAAGTT CTGTTGAACT AAAACACTTC TTCTTTGCTC
14701  AGGATGGCAA CGCTGCTATC AGTGATTATG ACTATTATCG TTATAATCTG
14751  CCAACAATGT GTGATATCAG ACAACTCCTA TTCGTAGTTG AAGTTGTTGA
14801  TAAATACTTT GATTGTTACG ATGGTGGCTG TATTAATGCC AACCAAGTAA
14851  TCGTTAACAA TCTGGATAAA TCAGCTGGTT TCCCATTTAA TAAATGGGGT
14901  AAGGCTAGAC TTTATTATGA CTCAATGAGT TATGAGGATC AAGATGCACT
14951  TTTCGCGTAT ACTAAGCGTA ATGTCATCCC TACTATAACT CAAATGAATC
15001  TTAAGTATGC CATTAGTGCA AAGAATAGAG CTCGCACCGT AGCTGGTGTC
15051  TCTATCTGTA GTACTATGAC AAATAGACAG TTTCATCAGA AATTATTGAA
15101  GTCAATAGCC GCCACTAGAG GAGCTACTGT GGTAATTGGA ACAAGCAAGT
15151  TTTACGGTGG CTGGCATAAT ATGTTAAAAA CTGTTTACAG TGATGTAGAA
15201  ACTCCACACC TTATGGGTTG GGATTATCCA AAATGTGACA GAGCCATGCC
15251  TAACATGCTT AGGATAATGG CCTCTCTTGT TCTTGCTCGC AAACATAACA
15301  CTTGCTGTAA CTTATCACAC CGTTTCTACA GGTTAGCTAA CGAGTGTGCG
15351  CAAGTATTAA GTGAGATGGT CATGTGTGGC GGCTCACTAT ATGTTAAACC
15401  AGGTGGAACA TCATCCGGTG ATGCTACAAC TGCTTATGCT AATAGTGTCT
```

Figure 65 (page 5 of 10)

Contig0001.TXT

```
15451  TTAACATTTG TCAAGCTGTT ACAGCCAATG TAAATGCACT TCTTTCAACT
15501  GATGGTAATA AGATAGCTGA CAAGTATGTC CGCAATCTAC AACACAGGCT
15551  CTATGAGTGT CTCTATAGAA ATAGGGATGT TGATCATGAA TTCGTGGATG
15601  AGTTTTACGC TTACCTGCGT AAACATTTCT CCATGATGAT TCTTTCTGAT
15651  GATGCCGTTG TGTGCTATAA CAGTAACTAT GCGGCTCAAG GTTTAGTAGC
15701  TAGCATTAAG AACTTTAAGG CAGTTCTTTA TTATCAAAAT AATGTGTTCA
15751  TGTCTGAGGC AAAATGTTGG ACTGAGACTG ACCTTACTAA AGGACCTCAC
15801  GAATTTTGCT CACAGCATAC AATGCTAGTT AAACAAGGAG ATGATTACGT
15851  GTACCTGCCT TACCCAGATC CATCAAGAAT ATTAGGCGCA GGCTGTTTTG
15901  TCGATGATAT TGTCAAAACA GATGGTACAC TTATGATTGA AAGGTTCGTG
15951  TCACTGGCTA TTGATGCTTA CCCACTTACA AAACATCCTA ATCAGGAGTA
16001  TGCTGATGTC TTTCACTTGT ATTTACAATA CATTAGAAAG TTACATGATG
16051  AGCTTACTGG CCACATGTTG GACATGTATT CCGTAATGCT AACTAATGAT
16101  AACACCTCAC GGTACTGGGA ACCTGAGTTT TATGAGGCTA TGTACACACC
16151  ACATACAGTC TTGCAGGCTG TAGGTGCTTG TGTATTGTGC AATTCACAGA
16201  CTTCACTTCG TTGCGGTGCC TGTATTAGGA GACCATTCCT ATGTTGCAAG
16251  TGCTGCTATG ACCATGTCAT TTCAACATCA CACAAATTAG TGTTGTCTGT
16301  TAATCCCTAT GTTTGCAATG CCCCAGGTTG TGATGTCACT GATGTGACAC
16351  AACTGTATCT AGGAGGTATG AGCTATTATT GCAAGTCACA TAAGCCTCCC
16401  ATTAGTTTTC CATTATGTGC TAATGGTCAG GTTTTTGGTT TATACAAAAA
16451  CACATGTGTA GGCAGTGACA ATGTCACTGA CTTCAATGCG ATAGCAACAT
16501  GTGATTGGAC TAATGCTGGC GATTACATAC TTGCCAACAC TTGTACTGAG
16551  AGACTCAAGC TTTTCGCAGC AGAAACGCTC AAAGCCACTG AGGAAACATT
16601  TAAGCTGTCA TATGGTATTG CTACTGTACG CGAAGTACTC TCTGACAGAG
16651  AATTGCATCT TTCATGGGAG GTTGGAAAAC CTAGACCACC ATTGAACAGA
16701  AACTATGTCT TTACTGGTTA CCGTGTAACT AAAAATAGTA AAGTACAGAT
16751  TGGAGAGTAC ACCTTTGAAA AAGGTGACTA TGGTGATGCT GTTGTGTACA
16801  GAGGTACTAC GACATACAAG TTGAATGTTG GTGATTACTT TGTGTTGACA
16851  TCTCACACTG TAATGCCACT TAGTGCACCT ACTCTAGTGC CACAAGAGCA
16901  CTATGTGAGA ATTACTGGCT TGTACCCAAC ACTCAACATC TCAGATGAGT
16951  TTTCTAGCAA TGTTGCAAAT TATCAAAAGG TCGGCATGCA AAAGTACTCT
17001  ACACTCCAAG GACCACCTGG TACTGGTAAG AGTCATTTTG CCATCGGACT
17051  TGCTCTCTAT TACCCATCTG CTCGCATAGT GTATACGGCA TGCTCTCATG
17101  CAGCTGTTGA TGCCCTATGT GAAAAGGCAT TAAAATATTT GCCCATAGAT
17151  AAATGTAGTA GAATCATACC TGCGCGTGCG CGCGTAGAGT GTTTTGATAA
17201  ATTCAAAGTG AATTCAACAC TAGAACAGTA TGTTTTCTGC ACTGTAAATG
17251  CATTGCCAGA AACAACTGCT GACATTGTAG TCTTTGATGA AATCTCTATG
17301  GCTACTAATT ATGACTTGAG TGTTGTCAAT GCTAGACTTC GTGCAAAACA
17351  CTACGTCTAT ATTGGCGATC CTGCTCAATT ACCAGCCCCC CGCACATTGC
17401  TGACTAAAGG CACACTAGAA CCAGAATATT TTAATTCAGT GTGCAGACTT
17451  ATGAAAACAA TAGGTCCAGA CATGTTCCTT GGAACTTGTC GCCGTTGTCC
17501  TGCTGAAATT GTTGACACTG TGAGTGCTTT AGTTTATGAC AATAAGCTAA
17551  AAGCACACAA GGATAAGTCA GCTCAATGCT TCAAAATGTT CTACAAAGGT
17601  GTTATTACAC ATGATGTTTC ATCTGCAATC AACAGACCTC AAATAGGCGT
17651  TGTAAGAGAA TTTCTTACAC GCAATCCTGC TTGGAGAAAA GCTGTTTTTA
17701  TCTCACCTTA TAATTCACAG AACGCTGTAG CTTCAAAAAT CTTAGGATTG
17751  CCTACGCAGA CTGTTGATTC ATCACAGGGT TCTGAATATG ACTATGTCAT
17801  ATTCACACAA ACTACTGAAA CAGCACACTC TTGTAATGTC AACCGCTTCA
17851  ATGTGGCTAT CACAAGGGCA AAAATTGGCA TTTTGTGCAT AATGTCTGAT
17901  AGAGATCTTT ATGACAAACT GCAATTTACA AGTCTAGAAA TACCACGTCG
17951  CAATGTGGCT ACATTACAAG CAGAAAATGT AACTGGACTT TTTAAGGACT
18001  GTAGTAAGAT CATTACTGGT CTTCATCCTA CACAGGCACC TACACACCTC
18051  AGCGTTGATA TAAAGTTCAA GACTGAAGGA TTATGTGTTG ACATACCAGG
18101  CATACCAAAG GACATGACCT ACCGTAGACT CATCTCTATG ATGGGTTTCA
18151  AAATGAATTA CCAAGTCAAT GGTTACCCTA ATATGTTTAT CACCCGCGAA
18201  GAAGCTATTC GTCACGTTCG TGCGTGGATT GGCTTTGATG TAGAGGGCTG
18251  TCATGCAACT AGAGATGCTG TGGGTACTAA CCTACCTCTC CAGCTAGGAT
18301  TTTCTACAGG TGTTAACTTA GTAGCTGTAC CGACTGGTTA TGTTGACACT
18351  GAAAATAACA CAGAATTCAC CAGAGTTAAT GCAAAACCTC CACCAGGTGA
18401  CCAGTTTAAA CATCTTATAC CACTCATGTA TAAAGGCTTG CCCTGGAATG
18451  TAGTGCGTAT TAAGATAGTA CAAATGCTCA GTGATACACT GAAAGGATTG
18501  TCAGACAGAG TCGTGTTCGT CCTTTGGGCG CATGGCTTTG AGCTTACATC
18551  AATGAAGTAC TTTGTCAAGA TTGGACCTGA AAGAACGTGT TGTCTGTGTG
```

Figure 65 (page 6 of 10)

Contig0001.TXT

```
18601  ACAAACGTGC AACTTGCTTT TCTACTTCAT CAGATACTTA TGCCTGCTGG
18651  AATCATTCTG TGGGTTTTGA CTATGTCTAT AACCCATTTA TGATTGATGT
18701  TCAGCAGTGG GGCTTTACGG GTAACCTTCA GAGTAACCAT GACCAACATT
18751  GCCAGGTACA TGGAAATGCA CATGTGGCTA GTTGTGATGC TATCATGACT
18801  AGATGTTTAG CAGTCCATGA GTGCTTTGTT AAGCGCGTTG ATTGGTCTGT
18851  TGAATACCCT ATTATAGGAG ATGAACTGAG GGTTAATTCT GCTTGCAGAA
18901  AAGTACAACA CATGGTTGTG AAGTCTGCAT TGCTTGCTGA TAAGTTTCCA
18951  GTTCTTCATG ACATTGGAAA TCCAAAGGCT ATCAAGTGTG TGCCTCAGGC
19001  TGAAGTAGAA TGGAAGTTCT ACGATGCTCA GCCATGTAGT GACAAAGCTT
19051  ACAAAATAGA GGAGCTCTTC TATTCTTATG CTACACATCA CGATAAATTC
19101  ACTGATGGTG TTTGTTTGTT TTGGAATTGT AACGTTGATC GTTACCCAGC
19151  CAATGCAATT GTGTGTAGGT TTGACACAAG AGCCTTGTCA AACTTGAACT
19201  TACCAGGCTG TGATGGTGGT AGTTTGTATG TGAATAAGCA TGCATTCCAC
19251  ACTCCAGCTT TCGATAAAAG TGCATTTACT AATTTAAAGC AATTGCCTTT
19301  CTTTTACTAT TCTGATAGTC CTTGTGAGTC TCATGGCAAA CAAGTAGTGT
19351  CGGATATTGA TTATGTTCCA CTCAAATCTG CTACGTGTAT TACACGATGC
19401  AATTTAGGTG GTGCTGTTTG CAGACACCAT GCAAATGAGT ACCGACAGTA
19451  CTTGGATGCA TATAATATGA TGATTTCTGC TGGATTTAGC CTATGGATTT
19501  ACAAACAATT TGATACTTAT AACCTGTGGA ATACATTTAC CAGGTTACAG
19551  AGTTTAGAAA ATGTGGCTTA TAATGTTGTT AATAAAGGAC ACTTTGATGG
19601  ACACGCCGGC GAAGCACCTG TTTCCATCAT TAATAATGCT GTTTACACAA
19651  AGGTAGATGG TATTGATGTG GAGATCTTTG AAAATAAGAC AACACTTCCT
19701  GTTAATGTTG CATTTGAGCT TTGGGCTAAG CGTAACATTA AACCAGTGCC
19751  AGAGATTAAG ATACTCAATA ATTTGGGTGT TGATATCGCT GCTAATACTG
19801  TAATCTGGGA CTACAAAAGA GAAGCCCCAG CACATGTATC TACAATAGGT
19851  GTCTGCACAA TGACTGACAT TGCCAAGAAA CCTACTGAGA GTGCTTGTTC
19901  TTCACTTACT GTCTTGTTTG ATGGTAGAGT GGAAGGACAG GTAGACCTTT
19951  TTAGAAACGC CCGTAATGGT GTTTTAATAA CAGAAGGTTC AGTCAAAGGT
20001  CTAACACCTT CAAAGGGACC AGCACAAGCT AGCGTCAATG GAGTCACATT
20051  AATTGGAGAA TCAGTAAAAA CACAGTTTAA CTACTTTAAG AAAGTAGACG
20101  GCATTATTCA ACAGTTGCCT GAAACCTACT TTACTCAGAG CAGAGACTTA
20151  GAGGATTTTA AGCCCAGATC ACAAATGGAA ACTGACTTTC TCGAGCTCGC
20201  TATGGATGAA TTCATACAGC GATATAAGCT CGAGGGCTAT GCCTTCGAAC
20251  ACATCGTTTA TGGAGATTTC AGTCATGGAC AACTTGGCGG TCTTCATTTA
20301  ATGATAGGCT TAGCCAAGCG CTCACAAGAT TCACCACTTA AATTAGAGGA
20351  TTTTATCCCT ATGGACAGCA CAGTGAAAAA TTACTTCATA ACAGATGCGC
20401  AAACAGGTTC ATCAAAATGT GTGTGTTCTG TGATTGATCT TTTACTTGAT
20451  GACTTTGTCG AGATAATAAA GTCACAAGAT TTGTCAGTGA TTTCAAAAGT
20501  GGTCAAGGTT ACAATTGACT ATGCTGAAAT TTCATTCATG CTTTGGTGTA
20551  AGGATGGACA TGTTGAAACC TTCTACCCAA AACTACAAGC AAGTCAAGCG
20601  TGGCAACCAG GTGTTGCGAT GCCTAACTTG TACAAGATGC AAAGAATGCT
20651  TCTTGAAAAG TGTGACCTTC AGAATTATGG TGAAAATGCT GTTATACCAA
20701  AAGGAATAAT GATGAATGTC GCAAAGTATA CTCAACTGTG TCAATACTTA
20751  AATACACTTA CTTTAGCTGT ACCCTACAAC ATGAGAGTTA TTCACTTTGG
20801  TGCTGGCTCT GATAAAGGAG TTGCACCAGG TACAGCTGTG CTCAGACAAT
20851  GGTTGCCAAC TGGCACACTA CTTGTCGATT CAGATCTTAA TGACTTCGTC
20901  TCCGACGCAG ATTCTACTTT AATTGGAGAC TGTGCAACAG TACATACGGC
20951  TAATAAATGG GACCTTATTA TTAGCGATAT GTATGACCCT AGGACCAAAC
21001  ATGTGACAAA AGAGAATGAC TCTAAAGAAG GGTTTTTCAC TTATCTGTGT
21051  GGATTTATAA AGCAAAAACT AGCCCTGGGT GGTTCTATAG CTGTAAAGAT
21101  AACAGAGCAT TCTTGGAATG CTGACCTTTA CAAGCTTATG GGCCATTTCT
21151  CATGGTGGAC AGCTTTTGTT ACAAATGTAA ATGCATCATC ATCGGAAGCA
21201  TTTTTAATTG GGGCTAACTA TCTTGGCAAG CCGAAGGAAC AAATTGATGG
21251  CTATACCATG CATGCTAACT ACATTTTCTG GAGGAACACA AATCCTATCC
21301  AGTTGTCTTC CTATTCACTC TTTGACATGA GCAAATTTCC TCTTAAATTA
21351  AGAGGAACTG CTGTAATGTC TCTTAAGGAG AATCAAATCA ATGATATGAT
21401  TTATTCTCTT CTGGAAAAAG GTAGGCTTAT CATTAGAGAA AACAACAGAG
21451  TTGTGGTTTC AAGTGATATT CTTGTTAACA ACTAAACGAA CATGTTTATT
21501  TTCTTATTAT TTCTTACTCT CACTAGTGGT AGTGACCTTG ACCGGTGCAC
21551  CACTTTTGAT GATGTTCAAG CTCCTAATTA CACTCAACAT ACTTCATCTA
21601  TGAGGGGGGT TTACTATCCT GATGAAATTT TTAGATCAGA CACTCTTTAT
21651  TTAACTCAGG ATTTATTTCT TCCATTTTAT TCTAATGTTA CAGGGTTTCA
21701  TACTATTAAT CATACGTTTG GCAACCCTGT CATACCTTTT AAGGATGGTA
```

Figure 65 (page 7 of 10)

Contig0001.TXT

```
21751  TTTATTTTGC TGCCACAGAG AAATCAAATG TTGTCCGTGG TTGGGTTTTT
21801  GGTTCTACCA TGAACAACAA GTCACAGTCG GTGATTATTA TTAACAATTC
21851  TACTAATGTT GTTATACGAG CATGTAACTT TGAATTGTGT GACAACCCTT
21901  TCTTTGCTGT TTCTAAACCC ATGGGTACAC AGACACATAC TATGATATTC
21951  GATAATGCAT TTAATTGCAC TTTCGAGTAC ATATCTGATG CCTTTTCGCT
22001  TGATGTTTCA GAAAAGTCAG GTAATTTTAA ACACTTACGA GAGTTTGTGT
22051  TTAAAAATAA AGATGGGTTT CTCTATGTTT ATAAGGGCTA TCAACCTATA
22101  GATGTAGTTC GTGATCTACC TTCTGGTTTT AACACTTTGA AACCTATTTT
22151  TAAGTTGCCT CTTGGTATTA ACATTACAAA TTTTAGAGCC ATTCTTACAG
22201  CCTTTTCACC TGCTCAAGAC ATTTGGGGCA CGTCAGCTGC AGCCTATTTT
22251  GTTGGCTATT TAAAGCCAAC TACATTTATG CTCAAGTATG ATGAAAATGG
22301  TACAATCACA GATGCTGTTG ATTGTTCTCA AAATCCACTT GCTGAACTCA
22351  AATGCTCTGT TAAGAGCTTT GAGATTGACA AAGGAATTTA CCAGACCTCT
22401  AATTTCAGGG TTGTTCCCTC AGGAGATGTT GTGAGATTCC CTAATATTAC
22451  AAACTTGTGT CCTTTTGGAG AGGTTTTTAA TGCTACTAAA TTCCCTTCTG
22501  TCTATGCATG GGAGAGAAAA AAAATTTCTA ATTGTGTTGC TGATTACTCT
22551  GTGCTCTACA ACTCAACATT TTTTTCAACC TTTAAGTGCT ATGGCGTTTC
22601  TGCCACTAAG TTGAATGATC TTTGCTTCTC CAATGTCTAT GCAGATTCTT
22651  TTGTAGTCAA GGGAGATGAT GTAAGACAAA TAGCGCCAGG ACAAACTGGT
22701  GTTATTGCTG ATTATAATTA TAAATTGCCA GATGATTTCA TGGGTTGTGT
22751  CCTTGCTTGG AATACTAGGA ACATTGATGC TACTTCAACT GGTAATTATA
22801  ATTATAAATA TAGGTATCTT AGACATGGCA AGCTTAGGCC CTTTGAGAGA
22851  GACATATCTA ATGTGCCTTT CTCCCCTGAT GGCAAACCTT GCACCCCACC
22901  TGCTCTTAAT TGTTATTGGC CATTAAATGA TTATGGTTTT TACACCACTA
22951  CTGGCATTGG CTACCAACCT TACAGAGTTG TAGTACTTTC TTTTGAACTT
23001  TTAAATGCAC CGGCCACGGT TTGTGGACCA AAATTATCCA CTGACCTTAT
23051  TAAGAACCAG TGTGTCAATT TTAATTTTAA TGGACTCACT GGTACTGGTG
23101  TGTTAACTCC TTCTTCAAAG AGATTTCAAC CATTTCAACA ATTTGGCCGT
23151  GATGTTTCTG ATTTCACTGA TTCCGTTCGA GATCCTAAAA CATCTGAAAT
23201  ATTAGACATT TCACCTTGCT CTTTTGGGGG TGTAAGTGTA ATTACACCTG
23251  GAACAAATGC TTCATCTGAA GTTGCTGTTC TATATCAAGA TGTTAACTGC
23301  ACTGATGTTT CTACAGCAAT TCATGCAGAT CAACTCACAC CAGCTTGGCG
23351  CATATATTCT ACTGGAAACA ATGTATTCCA GACTCAAGCA GGCTGTCTTA
23401  TAGGAGCTGA GCATGTCGAC ACTTCTTATG AGTGCGACAT TCCTATTGGA
23451  GCTGGCATTT GTGCTAGTTA CCATACAGTT TCTTTATTAC GTAGTACTAG
23501  CCAAAAATCT ATTGTGGCTT ATACTATGTC TTTAGGTGCT GATAGTTCAA
23551  TTGCTTACTC TAATAACACC ATTGCTATAC CTACTAACTT TTCAATTAGC
23601  ATTACTACAG AAGTAATGCC TGTTTCTATG GCTAAAACCT CCGTAGATTG
23651  TAATATGTAC ATCTGCGGAG ATTCTACTGA ATGTGCTAAT TTGCTTCTCC
23701  AATATGGTAG CTTTTGCACA CAACTAAATC GTGCACTCTC AGGTATTGCT
23751  GCTGAACAGG ATCGCAACAC ACGTGAAGTG TTCGCTCAAG TCAAACAAAT
23801  GTACAAAACC CCAACTTTGA AATATTTTGG TGGTTTTAAT TTTTCACAAA
23851  TATTACCTGA CCCTCTAAAG CCAACTAAGA GGTCTTTTAT TGAGGACTTG
23901  CTCTTTAATA AGGTGACACT CGCTGATGCT GGCTTCATGA AGCAATATGG
23951  CGAATGCCTA GGTGATATTA ATGCTAGAGA TCTCATTTGT GCGCAGAAGT
24001  TCAATGGACT TACAGTGTTG CCACCTCTGC TCACTGATGA TATGATTGCT
24051  GCCTACACTG CTGCTCTAGT TAGTGGTACT GCCACTGCTG GATGGACATT
24101  TGGTGCTGGC GCTGCTCTTC AAATACCTTT TGCTATGCAA ATGGCATATA
24151  GGTTCAATGG CATTGGAGTT ACCCAAAATG TTCTCTATGA GAACCAAAAA
24201  CAAATCGCCA ACCAATTTAA CAAGGCGATT AGTCAAATTC AAGAATCACT
24251  TACAACAACA TCAACTGCAT TGGGCAAGCT GCAAGACGTT GTTAACCAGA
24301  ATGCTCAAGC ATTAAACACA CTTGTTAAAC AACTTAGCTC TAATTTTGGT
24351  GCAATTTCAA GTGTGCTAAA TGATATCCTT TCGCGACTTG ATAAAGTCGA
24401  GGCGGAGGTA CAAATTGACA GGTTAATTAC AGGCAGACTT CAAAGCCTTC
24451  AAACCTATGT AACACAACAA CTAATCAGGG CTGCTGAAAT CAGGGCTTCT
24501  GCTAATCTTG CTGCTACTAA AATGTCTGAG TGTGTTCTTG GACAATCAAA
24551  AAGAGTTGAC TTTTGTGGAA AGGGCTACCA CCTTATGTCC TTCCCACAAG
24601  CAGCCCCGCA TGGTGTTGTC TTCCTACATG TCACGTATGT GCCATCCCAG
24651  GAGAGGAACT TCACCACAGC GCCAGCAATT TGTCATGAAG GCAAAGCATA
24701  CTTCCCTCGT GAAGGTGTTT TTGTGTTTAA TGGCACTTCT TGGTTTATTA
24751  CACAGAGGAA CTTCTTTTCT CCACAAATAA TTACTACAGA CAATACATTT
24801  GTCTCAGGAA ATTGTGATGT CGTTATTGGC ATCATTAACA ACACAGTTTA
24851  TGATCCTCTG CAACCTGAGC TCGACTCATT CAAAGAAGAG CTGGACAAGT
```

Figure 65 (page 8 of 10)

Contig0001.TXT

```
24901  ACTTCAAAAA TCATACATCA CCAGATGTTG ATCTTGGCGA CATTTCAGGC
24951  ATTAACGCTT CTGTCGTCAA CATTCAAAAA GAAATTGACC GCCTCAATGA
25001  GGTCGCTAAA AATTTAAATG AATCACTCAT TGACCTTCAA GAATTGGGAA
25051  AATATGAGCA ATATATTAAA TGGCCTTGGT ATGTTTGGCT CGGCTTCATT
25101  GCTGGACTAA TTGCCATCGT CATGGTTACA ATCTTGCTTT GTTGCATGAC
25151  TAGTTGTTGC AGTTGCCTCA AGGGTGCATG CTCTTGTGGT TCTTGCTGCA
25201  AGTTTGATGA GGATGACTCT GAGCCAGTTC TCAAGGGTGT CAAATTACAT
25251  TACACATAAA CGAACTTATG GATTTGTTTA TGAGATTTTT TACTCTTGGA
25301  TCAATTACTG CACAGCCAGT AAAAATTGAC AATGCTTCTC CTGCAAGTAC
25351  TGTTCATGCT ACAGCAACGA TACCGCTACA AGCCTCACTC CCTTTCGGAT
25401  GGCTTGTTAT TGGCGTTGCA TTTCTTGCTG TTTTTCAGAG CGCTACCAAA
25451  ATAATTGCGC TCAATAAAAG ATGGCAGCTA GCCCTTTATA AGGGCTTCCA
25501  GTTCATTTGC AATTTACTGC TGCTATTTGT TACCATCTAT TCACATCTTT
25551  TGCTTGTCGC TGCAGGTATG GAGGCGCAAT TTTTGTACCT CTATGCCTTG
25601  ATATATTTTC TACAATGCAT CAACGCATGT AGAATTATTA TGAGATGTTG
25651  GCTTTGTTGG AAGTGCAAAT CCAAGAACCC ATTACTTTAT GATGCCAACT
25701  ACTTTGTTTG CTGGCACACA CATAACTATG ACTACTGTAT ACCATATAAC
25751  AGTGTCACAG ATACAATTGT CGTTACTGAA GGTGACGGCA TTTCAACACC
25801  AAAACTCAAA GAAGACTACC AAATTGGTGG TTATTCTGAG GATAGGCACT
25851  CAGGTGTTAA AGACTATGTC GTTGTACATG GCTATTTCAC CGAAGTTTAC
25901  TACCAGCTTG AGTCTACACA AATTACTACA GACACTGGTA TTGAAAATGC
25951  TACATTCTTC ATCTTTAACA AGCTTGTTAA AGACCCACCG AATGTGCAAA
26001  TACACACAAT CGACGGCTCT TCAGGAGTTG CTAATCCAGC AATGGATCCA
26051  ATTTATGATG AGCCGACGAC GACTACTAGC GTGCCTTTGT AAGCACAAGA
26101  AAGTGAGTAC GAACTTATGT ACTCATTCGT TTCGGAAGAA ACAGGTACGT
26151  TAATAGTTAA TAGCGTACTT CTTTTTCTTG CTTTCGTGGT ATTCTTGCTA
26201  GTCACACTAG CCATCCTTAC TGCGCTTCGA TTGTGTGCGT ACTGCTGCAA
26251  TATTGTTAAC GTGAGTTTAG TAAAACCAAC GGTTTACGTC TACTCGCGTG
26301  TTAAAAATCT GAACTCTTCT GAAGGAGTTC CTGATCTTCT GGTCTAAACG
26351  AACTAACTAT TATTATTATT CTGTTTGGAA CTTTAACATT GCTTATCATG
26401  GCAGACAACG GTACTATTAC CGTTGAGGAG CTTAAACAAC TCCTGGAACA
26451  ATGGAACCTA GTAATAGGTT TCCTATTCCT AGCCTGGATT ATGTTACTAC
26501  AATTTGCCTA TTCTAATCGG AACAGGTTTT TGTACATAAT AAAGCTTGTT
26551  TTCCTCTGGC TCTTGTGGCC AGTAACACTT GCTTGTTTTG TGCTTGCTGC
26601  TGTCTACAGA ATTAATTGGG TGACTGGCGG GATTGCGATT GCAATGGCTT
26651  GTATTGTAGG CTTGATGTGG CTTAGCTACT TCGTTGCTTC CTTCAGGCTG
26701  TTTGCTCGTA CCCGCTCAAT GTGGTCATTC AACCCAGAAA CAAACATTCT
26751  TCTCAATGTG CCTCTCCGGG GGACAATTGT GACCAGACCG CTCATGGAAA
26801  GTGAACTTGT CATTGGTGCT GTGATCATTC GTGGTCACTT GCGAATGGCC
26851  GGACACCCCC TAGGGCGCTG TGACATTAAG GACCTGCCAA AAGAGATCAC
26901  TGTGGCTACA TCACGAACGC TTTCTTATTA CAAATTAGGA GCGTCGCAGC
26951  GTGTAGGCAC TGATTCAGGT TTTGCTGCAT ACAACCGCTA CCGTATTGGA
27001  AACTATAAAT TAAATACAGA CCACGCCGGT AGCAACGACA ATATTGCTTT
27051  GCTAGTACAG TAAGTGACAA CAGATGTTTC ATCTTGTTGA CTTCCAGGTT
27101  ACAATAGCAG AGATATTGAT TATCATTATG AGGACTTTCA GGATTGCTAT
27151  TTGGAATCTT GACGTTATAA TAAGTTCAAT AGTGAGACAA TTATTTAAGC
27201  CTCTAACTAA GAAGAATTAT TCGGAGTTAG ATGATGAAGA ACCTATGGAG
27251  TTAGATTATC CATAAAACGA ACATGAAAAT TATTCTCTTC CTGACATTGA
27301  TTGTATTTAC ATCTTGCGAG CTATATCACT ATCAGGAGTG TGTTAGAGGT
27351  ACGACTGTAC TACTAAAAGA ACCTTGCCCA TCAGGAACAT ACGAGGGCAA
27401  TTCACCATTT CACCCTCTTG CTGACAATAA ATTTGCACTA ACTTGCACTA
27451  GCACACACTT TGCTTTTGCT TGTGCTGACG GTACTCGACA TACCTATCAG
27501  CTGCGTGCAA GATCAGTTTC ACCAAAACTT TTCATCAGAC AAGAGGAGGT
27551  TCAACAAGAG CTCTACTCGC CACTTTTTCT CATTGTTGCT GCTCTAGTAT
27601  TTTTAATACT TTGCTTCACC ATTAAGAGAA AGACAGAATG AATGAGCTCA
27651  CTTTAATTGA CTTCTATTTG TGCTTTTTAG CCTTTCTGCT ATTCCTTGTT
27701  TTAATAATGC TTATTATATT TTGGTTTTCA CTCGAAATCC AGGATCTAGA
27751  AGAACCTTGT ACCAAAGTCT AAACGAACAT GAAACTTCTC ATTGTTTTGA
27801  CTTGTATTTC TCTATGCAGT TGCATATGCA CTGTAGTACA GCGCTGTGCA
27851  TCTAATAAAC CTCATGTGCT TGAAGATCCT TGTAAGGTAC AACACTAGGG
27901  GTAATACTTA TAGCACTGCT TGGCTTTGTG CTCTAGGAAA GGTTTTACCT
27951  TTTCATAGAT GGCACACTAT GGTTCAAACA TGCACACCTA ATGTTACTAT
28001  CAACTGTCAA GATCCAGCTG GTGGTGCGCT TATAGCTAGG TGTTGGTACC
```

Figure 65 (page 9 of 10)

Contig0001.TXT

```
28051  TTCATGAAGG TCACCAAACT GCTGCATTTA GAGACGTACT TGTTGTTTTA
28101  AATAAACGAA CAAATTAAAA TGTCTGATAA TGGACCCCAA TCAAACCAAC
28151  GTAGTGCCCC CCGCATTACA TTTGGTGGAC CCACAGATTC AACTGACAAT
28201  AACCAGAATG GAGGACGCAA TGGGGCAAGG CCAAAACAGC GCCGACCCCA
28251  AGGTTTACCC AATAATACTG CGTCTTGGTT CACAGCTCTC ACTCAGCATG
28301  GCAAGGAGGA ACTTAGATTC CCTCGAGGCC AGGGCGTTCC AATCAACACC
28351  AATAGTGGTC CAGATGACCA AATTGGCTAC TACCGAAGAG CTACCCGACG
28401  AGTTCGTGGT GGTGACGGCA AAATGAAAGA GCTCAGCCCC AGATGGTACT
28451  TCTATTACCT AGGAACTGGC CCAGAAGCTT CACTTCCCTA CGGCGCTAAC
28501  AAAGAAGGCA TCGTATGGGT TGCAACTGAG GGAGCCTTGA ATACACCCAA
28551  AGACCACATT GGCACCCGCA ATCCTAATAA CAATGCTGCC ACCGTGCTAC
28601  AACTTCCTCA AGGAACAACA TTGCCAAAAG GCTTCTACGC AGAGGGAAGC
28651  AGAGGCGGCA GTCAAGCCTC TTCTCGCTCC TCATCACGTA GTCGCGGTAA
28701  TTCAAGAAAT TCAACTCCTG GCAGCAGTAG GGGAAATTCT CCTGCTCGAA
28751  TGGCTAGCGG AGGTGGTGAA ACTGCCCTCG CGCTATTGCT GCTAGACAGA
28801  TTGAACCAGC TTGAGAGCAA AGTTTCTGGT AAAGGCCAAC AACAACAAGG
28851  CCAAACTGTC ACTAAGAAAT CTGCTGCTGA GGCATCTAAA AAGCCTCGCC
28901  AAAAACGTAC TGCCACAAAA CAGTACAACG TCACTCAAGC ATTTGGGAGA
28951  CGTGGTCCAG AACAAACCCA AGGAAATTTC GGGGACCAAG ACCTAATCAG
29001  ACAAGGAACT GATTACAAAC ATTGGCCGCA AATTGCACAA TTTGCTCCAA
29051  GTGCCTCTGC ATTCTTTGGA ATGTCACGCA TTGGCATGGA AGTCACACCT
29101  TCGGGAACAT GGCTGACTTA TCATGGAGCC ATTAAATTGG ATGACAAAGA
29151  TCCACAATTC AAAGACAACG TCATACTGCT GAACAAGCAC ATTGACGCAT
29201  ACAAAACATT CCCACCAACA GAGCCTAAAA AGGACAAAAA GAAAAAGACT
29251  GATGAAGCTC AGCCTTTGCC GCAGAGACAA AAGAAGCAGC CCACTGTGAC
29301  TCTTCTTCCT GCGGCTGACA TGGATGATTT CTCCAGACAA CTTCAAAATT
29351  CCATGAGTGG AGCTTCTGCT GATTCAACTC AGGCATAAAC ACTCATGATG
29401  ACCACACAAG GCAGATGGGC TATGTAAACG TTTTCGCAAT TCCGTTTACG
29451  ATACATAGTC TACTCTTGTG CAGAATGAAT TCTCGTAACT AAACAGCACA
29501  AGTAGGTTTA GTTAACTTTA ATCTCACATA GCAATCTTTA ATCAATGTGT
29551  AACATTAGGG AGGACTTGAA AGAGCCACCA CATTTTCATC GAGGCCACGC
29601  GGAGTACGAT CGAGGGTACA GTGAATAATG CTAGGGAGAG CTGCCTATAT
29651  GGAAGAGCCC TAATGTGTAA AATTAATTTT AGTAGTGCTA TCCCCATGTG
29701  ATTTTAATAG CTTCTTAGGA GAATGAC
```

Figure 65 (page 10 of 10)